Immunophilins: An Investigation into Function and Structure

A Thesis

Submitted for the Degree of

Doctor of Philosophy

by

Jacqueline Dornan, B.Sc.

Structural Biochemistry Group

The Institute of Cell and Molecular Biology

University of Edinburgh

September 2001

SUMMARY OF CONTENTS

ABSTRACT

The Immunophilin family. The immunophilin family of proteins consists of two members, Cyclophilins and FK-506 binding proteins (FKBP's). Immunophilins are ubiquitous and highly evolutionarily conserved. Initially characterised by their high affinity binding to immunosuppressant drugs Cyclosporin A (CsA) and FK-506 respectively, these proteins were shown to have intrinsic enzymatic activity, exhibiting catalysis of peptidyl prolyl cis-trans isomerisation (PPIase activity) around an Xaa-Pro bond. Subsequent studies have shown the involvement of immunophilins in many cellular processes, for example as components of the multi-protein complex responsible for the assembly and maturation of steroid hormone receptors.

The work described in this thesis focused on the over-expression, purification, biochemical characterisation, crystallisation and 3-D structure determination of three members of the immunophilin family. The overall goal of the project was to use biochemical and structural information derived from these studies to assist in the elucidation of the various roles played by these proteins in diverse systems.

Cyclophilin 3. Cloning and expression of eleven cyclophilin homologues from the free-living nematode *Caenorhabditis elegans* (*C. elegans*) has recently been reported. Cyclophilin 3 (Cyp-3), one of the most abundantly expressed isoforms, was chosen as the initial target for further study. The structure of the protein was solved to high resolution, allowing identification of a number of key structural features that differ from those of the archetypal cyclophilin, Cyclophilin A. These newly determined features, which may prove to be functionally significant, were used to define a new cyclophilin subfamily. The PPIase activity and degree of inhibition of the protein by CsA were also characterised. A series of experiments involving metal binding to Cyp-3 were also carried out. Results presented include binding of an anti-arthritic gold complex to Cyp-3. Attempts to induce metal catalysed oxidation of cysteines present in the reduced form in the structure of Cyp-3 are described. The first reported direct determination of a ligand binding constant in protein crystals is

reported, following correlation of data obtained from both solution and X-ray crystallographic studies of the binding of a di-peptide ligand (Ala-Pro) to Cyp-3.

Cyclophilin 40. Cyclophilin 40 (Cyp 40) is a two domain immunophilin containing a conserved cyclophilin domain linked *via* a charged linker region to the C-terminal domain comprised largely of three copies of the tetratricopeptide repeat (TPR) motif. Cyp 40 constitutes part of the "mature" steroid hormone receptor complex, interactions with Hsp90 have been described, although the exact nature of the interaction has not been well defined. Using cloned bovine Cyp 40, two very different crystal forms were grown. The resultant structures, monoclinic and tetragonal Cyp 40, are illustrated and described. The structure of monoclinic Cyp 40 obtained was used as the basis for comparative protein modelling studies on a number of cyclophilin 40 type proteins from different species, the results of which are included. The second major objective was to use the experimentally determined protein structure to derive new insights into the interactions seen with an important partner protein, Hsp90. A model of the terminal five residues of Hsp90 (MEEVD) binding to Cyp 40 is reported. A modelled interaction between 95 C-terminal residues of hsp90 and the groove formed by the folding of the TPR domain of Cyp 40 is also described. The over-expression, purification and crystallisation trials of two Hsp90 C-terminal fragments are also reported.

FKBP 22. Isolation of FKBP 22 from the endoplasmic reticulum of N*eurospora crassa* (*N. crassa*) has recently been reported. The C-terminal sequence appears to be unique among other FKBP sequences, being rather highly charged and is consistent with an amphipathic helical conformation. Crystallisation and preliminary partial X-ray structure determination of FKBP22 from *N. crassa* are reported.

Declaration

The work presented in this thesis is the original work of the author, any collaboration has been indicated clearly. This thesis has been composed by the author and has not been submitted in whole or in part for any other degree. Some of the results have been published.

Jacqueline Dornan

Dedicated to my parents,

James and Sheila Dornan

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Malcolm D. Walkinshaw, firstly for the opportunity to register for a PhD and also for his invaluable advice, enthusiasm and encouragement throughout the duration of the project and beyond.

I would also like to thank Dr. Paul Taylor for all his help, advice and patience in dealing with crystallographic and computational problems, and for his friendship.

I am also indebted to all members of the Walkinshaw group, both past and present, for helping me with various aspects of the projects described in this thesis. In particular I would like to thank Dr. George Kontopidis for his help with the PPIase assay and Dr. Holger Husi for advice on protein purification during the early stages of the project.

Thanks are also due to many people including Violet Anderson, Sandra Bruce, Dr. Nichola Picken, Mary Russell, Dr. Martin Wear and Dr. Su-ying Wu, for their friendship and support throughout good times and bad.
Members of the Structural Biology Group are also due thanks for making work an entertaining place to be and writing-up bearable.

I would also like to acknowledge Dr. Tony Page, Dr. Tom Ratajczak and Professor Maximillian Tropschug for providing the clones used throughout the duration of this project, and for their continued interest in the results.

APB	Amersham Pharmacia Biotech
AS	Ammonium sulphate
β-Me	2 Mercaptoethanol
CsA	Cyclosporin A
Cyp	Cyclophilin
Cyp A	Human cyclophilin A protein
DMSO	Dimethyl sulphoxide
DTNB	5,5'-Dithiobis(2-nitrobenzoic acid)
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FKBP	FK-506 binding protein
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GR	Glucocorticoid receptor
GST	Glutathione S transferase
HEPES	4 (2-Hydroxyethyl) piperazine-1 ethane sulphonic acid
HLA	Human leukocyte antigens
Hsp	Heat shock protein
ICMB	Institute of Cell and Molecular Biology
IPTG	Isopropyl β-D- thiogalactopyranoside
kDa	Kilodalton
LB	Luria Bertani media
LC-ES-MS	Liquid chromatography electrospray mass spectrometry
MES	2-(N-Morpholino) ethanesulphonic acid
MPEG	Methoxypolyethyleneglycol
MWT	Molecular weight
NF-AT	Nuclear factor of activated T cells
NMR	Nuclear magnetic resonance
OD_{600}	Optical density at 600nm
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PEG	Polyethyleneglycol
pET	Plasmid under the expression of the T7 promoter
pI	Isoelectric point
pNA	pNitroaniline
PPIase	Peptidyl prolyl *cis-trans* isomerase
rmsd	Root mean square deviation
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase polymerase chain reaction
RyR	Ryanodine receptor
SDS	Sodium dodecyl sulphate
SHR	Steroid hormone receptor
SN	Supernatant
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TPR	Tetratricopeptide

TABLE OF CONTENTS

CHAPTER 1.
INTRODUCTION

CHAPTER 3.

PURIFICATION, CHARACTERISATION AND CRYSTALLISATION OF CYCLOPHILIN 3 FROM THE FREE LIVING NEMATODE CAENORHABDITIS ELEGANS

CHAPTER 3B

FURTHER STRUCTURAL AND ENZYMATIC STUDIES OF CYCLOPHILIN 3

CHAPTER 4

PURIFICATION, CHARACTERISATION AND CRYSTALLISATION OF BOVINE CYCLOPHILIN 40

CHAPTER 5

PURIFICATION AND CHARACTERISATION OF C-TERMINAL FRAGMENTS OF HEAT SHOCK PROTEIN 90

CHAPTER 6
MODELLING STUDIES

CHAPTER 6B

MODELLED INTERACTIONS BETWEEN THE C-TERMINAL DOMAINS OF HSP90 AND CYP 40

CHAPTER 7.

CRYSTALLISATION AND PRELIMINARY PARTIAL X-RAY STRUCTURE DETERMINATION OF FKBP 22 FROM NEUROSPORA CRASSA

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1

THE IMMUNOPHILIN FAMILY

This chapter will introduce the Immunophilin family of proteins to be discussed in this thesis. Background information on the isolation, characterisation and activities of this family of proteins will be described.

1.1.1

The immunophilin super-family comprises two members, Cyclophilins and FK-506 Binding Proteins (FKBP's).

Immunophilins are ubiquitous and highly conserved proteins originally identified by their high affinity binding to immunosuppressant drugs. They can be divided into two classes, Cyclophilins bind Cyclosporin A, the FKBP's are binding proteins for the compounds FK-506 and rapamycin. The isolation, characterisation and structure determination of the archetypal members of these protein families will be described in greater detail later in this chapter.

Previous work demonstrating the immunophilin family served as ligands for immunosuppressant drugs greatly advanced studies of the mechanism of immunosuppression, but gave few clues as to the normal function of these proteins. Their degree of ubiquity however suggests additional roles in normal cellular processes, in the absence of exogenous drug.

The mechanism of immunosuppression involving a complex of the drugs (CsA and FK-506/Rapamycin) with their cognate immunophilins interacting with target proteins will be briefly described later in this chapter.

Immunophilins are also classified as Peptidyl Prolyl Isomerases, due to their intrinsic enzymatic activity. All members of the immunophilin family show peptidyl-prolyl isomerase (PPIase) activity. A third type of PPIase was discovered in 1990 and was assigned to a new protein family, the Parvulins. Parvulins however, do not appear to be associated with any known immunosuppressant drugs and therefore fall out-with the remit of this thesis.

1.1.2

Peptidyl-prolyl cis-trans isomerase activity (PPIase activity)

As previously described, three unrelated protein families Cyclophilins, FKBP's and Parvulins possess PPIase activity. In this section PPIase activity will be considered in more detail, and will include

- A brief discussion of the nature and occurrence of X-Pro peptide bonds
- The catalysis of protein folding brought about by prolyl isomerases
- Proposed PPIase reaction mechanisms
- Development of an assay for PPIase activity
- Examples of the efficiency of these enzymes

1.1.2.1

X-Pro peptide bonds

Planar peptide bonds in proteins have been shown to occur predominantly in the *trans* conformation. In 1990, Stewart *et al* investigated the occurrence of the *cis* isomer in protein structures. To do this they examined high resolution X-ray structures deposited with the Protein Data Bank. This study revealed there were very few *cis* peptide bonds present in native folded proteins. The exception was proline as approximately 10% of peptidyl-prolyl imide (X-Pro) bonds occur as *cis* isomers in native, folded proteins (Stewart et al., 1990). The importance of proline residues in defining and maintaining the structural integrity of proteins was described in a publication by Macarthur and Thornton (Macarthur and Thornton, 1991).

1.1.2.2

Catalysis of protein folding by prolyl isomerases and the role of PPIase activity in protein folding

Cyclophilins were initially characterised by their enzymatic activity. These proteins exhibited the ability to catalyse the slow *cis-trans* isomerisation step of proline-peptide (Xaa-Pro) bonds in the folding of oligopeptides and also accelerated the slow, rate limiting steps in the folding of several proteins (Fischer et al., 1984; Bachinger, 1987; Lang et al., 1987a).

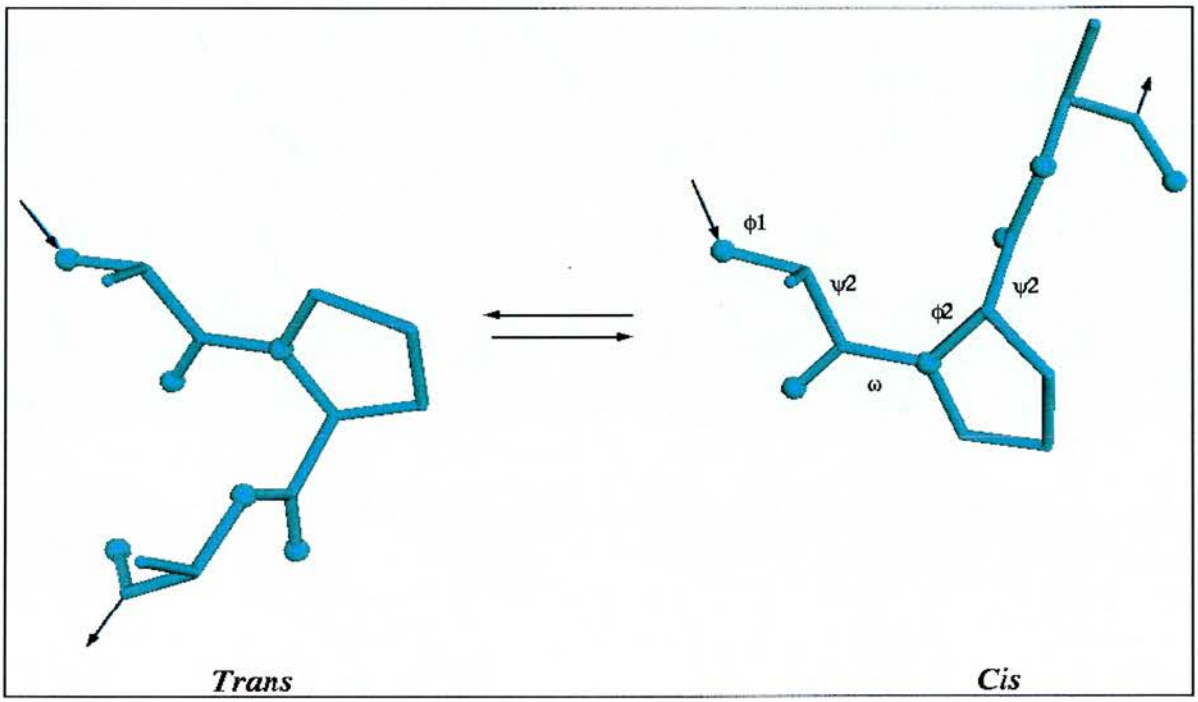


Figure 1

Cis-trans equilibrium and PPIase activity.

The conversion of *trans* proline to *cis* proline involves a rotation about the *imide* bond. PPIase activity reduces the energy barrier for the above reaction. Note the direction of the main chain shown by the arrows is opposite in the two conformers.

1.1.2.3

Proposed PPIase reaction mechanism

A number of different reaction mechanisms have been proposed to explain the catalysis of *cis-trans* isomerisation induced by immunophilins. These have included:

1. The formation of a tetrahedral intermediate formed by the nucleophilic addition of a cysteine or nucleophilic residue to the carbonyl carbon of the Xaa-Pro bond (Fischer et al., 1989a).
2. Catalysis by protonation, involving side chain residues protonating the C-N peptide bond (Fischer et al., 1989a; Kofron et al., 1991).
3. Catalysis by desolvation which suggested desolvation of the substrate is the first step of the reaction, evidenced by the rate of *cis-trans* isomerisation being accelerated in nonpolar solvents (Radzicka et al., 1992; Eberhardt et al., 1992)
4. Catalysis by distortion/twisted amide model. It was proposed that the C-N bond was rotated allowing stabilisation of this distorted intermediate in the binding site of the molecule (Harrison and Stein, 1990; Rosen et al., 1990). These models arose from the suggestion that drugs inhibit the PPIase activity of the molecule by mimicking an intermediate state in the *cis-trans* interconversion of peptide and protein substrates. This appears to be currently the most favoured mechanism.
5. Solvent assisted mechanism. Similar to catalysis by distortion, except water molecues not binding site residues in the protein important in the mechanism, causing distortion of the amide bond (Ke et al., 1993; Zhao and Ke, 1996).

A summary of the various theories of the PPIase mechanism was provided in two review articles (Galat and Metcalfe, 1995; Ivery, 2000). The more recent review additionally discusses information on the potential PPIase mechanism provided from recent structural studies of a number of members of the immunophilin family solved with ligands bound in the active site.

The tight binding of CsA, FK-506 and rapamycin to their respective partners may be as a consequence of their ability to mimic the conformation of the transition state of the *cis-trans* reaction seen with peptide and protein substrates. Closer examination of protein structures with drugs bound has help shed more light onto this topic. A recent

paper (Ivery, 2000) provides a brief but reasonably comprehensive review of papers reporting both theoretical and mechanistic studies of PPIase activity.

1.1.2.4

Assays of PPIase activity

An assay for PPIase activity was developed by Fischer et al (Fischer et al., 1984). This spectrophotometric assay utilised the high degree of conformational selectivity of chymotrypsin towards peptide substrates of the type X-Pro-Phe-pNA. Chymotrypsin hydrolysis of the *p*-nitroanilide bond was shown to only occur if the prolyl amide bond was in the trans conformation, resulting in the release of chromogenic dye (pNA), accumulation of which could be measured spectrophotometrically. The ability of immunophilins to catalyse this *cis-trans* isomerisation therefore manifests itself in this particular assay by making more substrate available for hydrolysis with the consequent increase in measurable release of pNA.

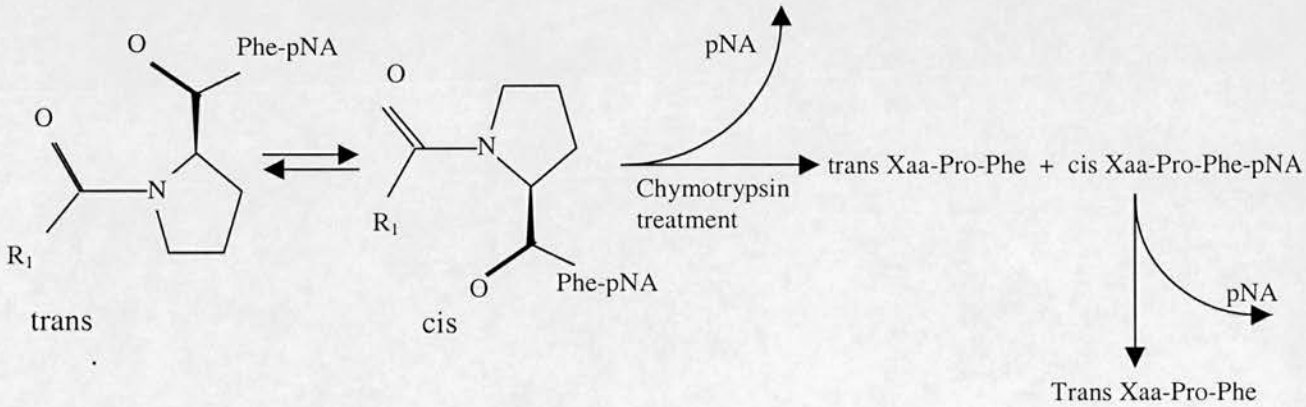


Figure 2

Schematic illustration of the PPIase assay (adapted from Goethel and Marahiel, 1999).

A number of competing processes exist within this assay, including the underlying, parallel, reversible *cis-trans* isomerisation and the irreversible chymotrypsin catalysed hydrolysis of the *trans* isomer conformation. In aqueous solutions, the peptide substrate is present largely in the *trans* confomer, resulting in hydrolysis with a very rapid burst phase with a small contribution from substrate isomerised during the experiment. This resulted in a consequently poor signal to noise ratio and was a limiting factor in the degree of usefulness of the assay as originally described. A modified version of the assay with a rigorous mathematical treatment of the resultant spectrophotometric progress curves was reported in 1991 (Kofron et al., 1991). The principal modifications to the experimental set-up were twofold.

(1) Use of a novel solvent system for the substrates (lithium chloride and trichloroethanol), to push the equilibrium towards the *cis* X-Pro isomer, resulting in an increase in available substrate from approximately 10% to 50 - 70%.

(2) Carrying out the assay at low temperature and also at a variety of wavelengths.

The improved assay allowed the study of the kinetics of the reaction under saturating conditions, allowing determination of both steady state parameters for PPIase substrates and inhibition patterns for a number of ligands.

1.1.2.5
Efficiency of PPIase activity

Modification to the PPIase activity assay as suggested by Kofron *et al* allowed determination of kinetic parameters of the catalytic efficiency of immunophilins. Cyp A and FKBP12 were shown to be efficient catalysts of isomerisation, (Kofron et al., 1991). Many other members of the family have been characterised, a number of examples are included in the table.

Protein	Catalytic efficiency k_{cat}/K_m	Reference
Cyp A	$14.6 \times 10^6 M^{-1}s^{-1}$	(Kofron et al., 1991)
BmCyp-1	$7.9 \times 10^6 M^{-1}s^{-1}$	(Page et al., 1995a)
Cyp 3	$2.4 \times 10^6 M^{-1}s^{-1}$	(Dornan et al., 1999a)
Cyp 40	$1.9 \times 10^6 M^{-1}s^{-1}$	(Kieffer et al., 1992)
FKBP 12	$0.66 \times 10^6 M^{-1}s^{-1}$	(Kofron et al., 1991)
FKBP 22	$6.94 \times 10^5 M^{-1}s^{-1}$	(Solscheid and Tropschug, 2000)
FKBP 52	$2.6 \times 10^4 M^{-1}s^{-1}$	(Peattie et al., 1992)

Table 1

Examples of the catalytic efficiency (expressed as k_{cat}/K_m) of selected immunophilins.

1.1.3

Mechanism resulting in immunosuppression

Elucidation of the cellular mechanisms underlying the immunosuppressive effects of CsA and FK-506 were the focus of a huge research effort for a number of years. The PPIase activity of these two proteins was shown to be non-essential for the resulting immunosuppressive activity, active site mutants of Cyp A were still able to bind CsA and calcineurin whilst having greatly reduced PPIase activity (Zydowsky et al., 1992).

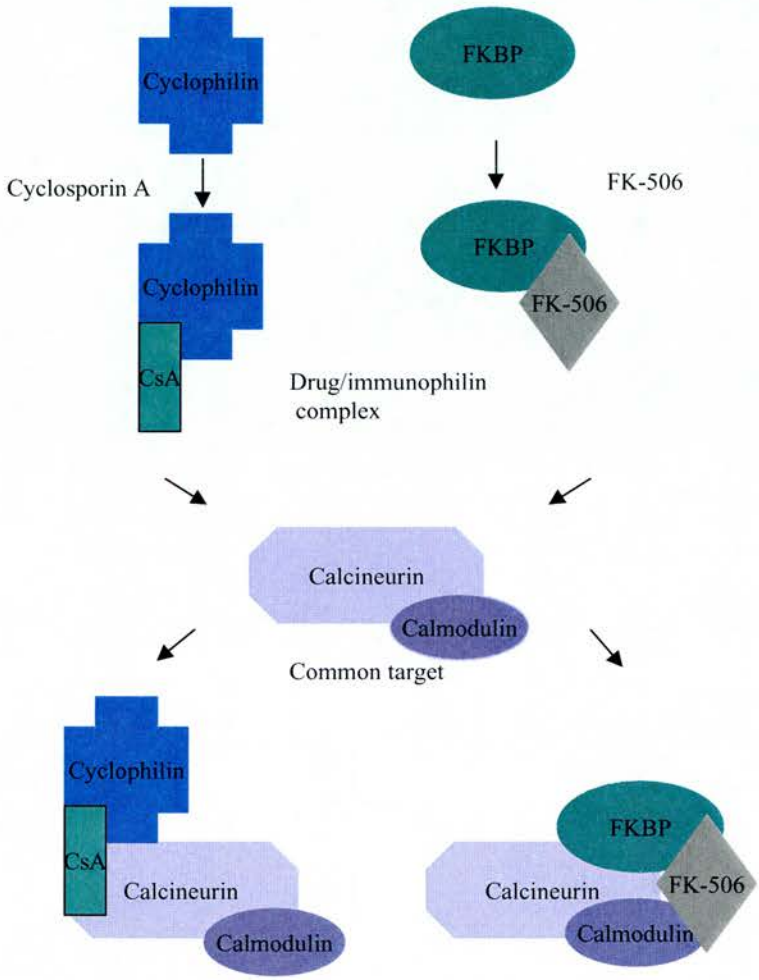


Figure 3
Mechanism resulting in immunosuppression.

Cyclophilin A/ FKBP12 form drug-immunophilin complexes with CsA and FK-506 respectively. These protein/drug complexes act directly upon the calmodulin dependent phosphatase calcineurin. The complexes of these two compounds were shown to inhibit calcium dependent signalling required for T-cell activation, by binding directly to calcineurin, thereby blocking nuclear transport of cytoplasmic transcription factors, primarily NF-AT, resulting in inhibition of activation of cytokine transcription (Liu et al., 1991a).

1.1.4

Roles of immunophilins in the absence of exogenous drugs

Despite having similar roles in immunosuppression and sharing PPIase activity, these two protein families are unrelated both in primary sequence and structure. Immunophilins have a broad phylogenetic distribution, (for reviews see (Galat and Metcalfe, 1995; Kay, 1996; Hamilton and Steiner, 1998), many species possessing numerous immunophilin isoforms. The roles played by immunophilins within the cell remains largely to be elucidated, however their degree of evolutionary conservation and ubiquity suggests an important role for these proteins. Evidence in support of a role for immunophilins acting in many pathways i.e. protein folding, regulation of cellular processes and molecular chaperoning of partner proteins has grown over recent years. A number of these potential roles will be discussed in later chapters. The focus will be in particular on the larger two domain immunophilin Cyclophilin 40.

The broad goal of the project to be discussed in this thesis was to use protein structures, in an attempt to shed light on the natural roles of selected members of the immunophilin family. The study of this intriguing family of proteins may provide insights into possible applications in a number of fields including development of targeted anti-parasitic drugs, protein-protein interactions and protein folding.

1.2

CYCLOPHILINS

1.2.1

Cyclosporin A, a potent and specific inhibitor of HLA mismatched allograft rejection

Cyclosporin A (CsA) was originally identified in the early seventies by Sandoz (now Novartis) during screening trials for natural microbial products which showed anti-bacterial and anti-fungal properties. It was discovered that a compound produced by the fungus *Tolypocladium inflatum Gams* (Borel et al., 1989) acted as an anti-fungal antibiotic. This compound was subsequently shown to possess not only antibiotic properties but also potent immunosuppressive activity, demonstrated by results of screening studies carried out in animals treated with purified fungal extracts. Extracts producing effects in animal models were further characterised and shown to consist mainly of CsA. CsA is a cyclic undecapeptide, its structure is shown in Figure 4.

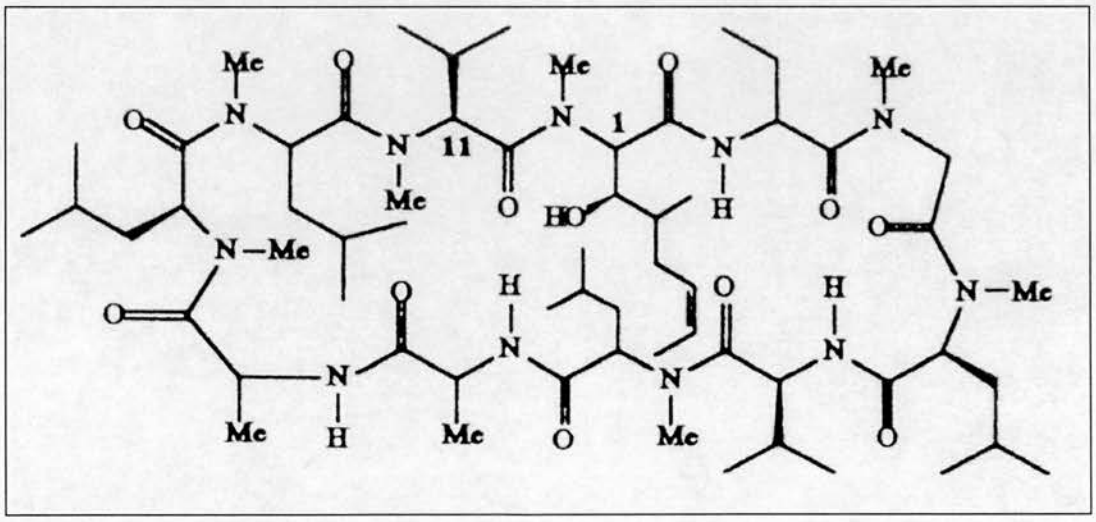


Figure 4

Chemical structure of Cyclosporin A, a natural cyclic undecapeptide from *Tolypocladium inflatum Gams.*

Following extensive trials, CsA received approval for clinical use in 1983, its principal application being to help prevent organ rejection following transplant surgery. CsA therapy has been credited with revolutionising the field of organ transplantation, increasing life expectancy and quality of life for many patients following such surgery. A number of additional uses for this drug have been found, however it was in the field of immunosuppression upon which most research interest focused.

The mechanism of action of this drug resulting in the immunosuppression seen in patients was largely unknown, therefore many groups were interested in isolating proteins interacting physiologically with CsA, in an attempt to determine its mode of action.

1.2.2

The isolation and characterisation of Cyclophilin A (Cyp A)

The first immunophilin to be discovered was the abundant cytosolic protein cyclophilin A (Cyp A), isolated by its high affinity binding to CsA. Cyp A was initially identified from a preparation of bovine thymocytes applied to a CsA affinity column, proteins which bound tightly were selected as targets for further study (Handschumacher et al., 1984). This group went on to purify to homogeneity a protein of approximately 17kDa, which showed the required high affinity binding to CsA. Material was isolated from two primary sources, bovine thymus cytosol and human spleen (Harding et al., 1986). Using this purified material Harding *et al* were able to elucidate the sequence of these two proteins, and show that they were essentially identical. The protein was called Cyclophilin A due to its association with Cyclosporin A.

1.2.3

Cyclophilin A and Peptidyl- prolyl cis-trans isomerase activity

In the early eighties a significant amount of research focused on a number of problems associated with the fast and slow phases of protein folding/refolding. A

particular area of interest was in the rate of *cis-trans* interconversion of proline containing peptides, with an emphasis on isolating and characterising proteins that were catalytically active. Numerous tissue extracts were investigated, amongst those showing demonstrable catalytic activity of the *cis-trans* interconversion was an extract prepared from porcine kidney cortex. The protein fraction responsible for the activity was purified and characterised, and classified enzymatically as a peptidyl-prolyl *cis-trans* isomerase (PPIase). N-terminal sequencing of the purified protein indicated that it was identical in sequence to the first 38 amino acids of Cyclophilin A, (Fischer et al., 1989b; Takahashi et al., 1989).

1.3
FK-506 BINDING PROTEINS

1.3.1
FK-506 and Rapamycin: immunosuppressive agents structurally unrelated to CsA

As a result of the success of CsA the search was on for new compounds with both more potent immunosuppressive activity and fewer side effects than CsA. This led to the discovery of two compounds, FK-506 isolated from the actinomycete *Streptomyces tsukubaensis*, (shown to be 50 times more potent an immunosuppressive agent than CsA) (Tanaka et al., 1987; Kino et al., 1987a; Kino et al., 1987b) and Rapamycin, another macrolide also first isolated as an anti-fungal agent from cultures of *Streptomyces hygroscopius* (Sehgal, 1995).

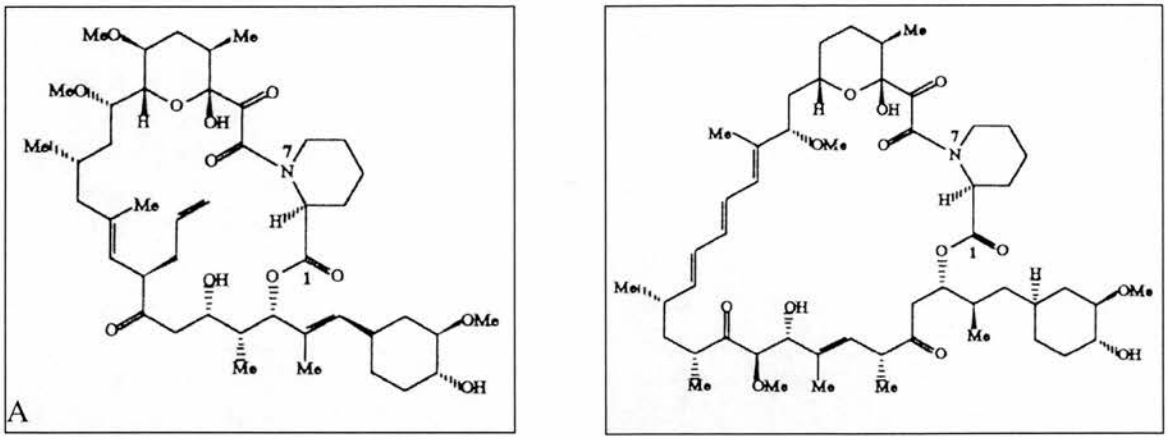


Figure 5

Chemical structures of (A) FK-506 from *Streptomyces tsukubaensis* and (B) Rapamycin from *Streptomyces hygroscopius*

The mechanisms of immunosuppression by these structurally unrelated compounds (CsA and FK-506) was shown to be remarkably similar, acting upon the common target molecule calcineurin as described earlier in this chapter. Rapamycin has an alternative, calcium independent, mode of action resulting in immunosuppression (Bierer et al., 1990; Morris, 1995).

1.3.2

The isolation and characterisation of FKBP's

Using an FK-506 affinity matrix, two groups independently isolated proteins from a number of sources, the JURKAT T cell line (Siekierka et al., 1989) and bovine thymus and human spleen (Harding et al., 1989). Due to the high degree of functional similarity with cyclophilins, these proteins were tested for PPIase activity. The authors proved the proteins did indeed possess PPIase activity and therefore belonged to the immunophilin family.

Cloning and expression of FKBP 12

This small, abundant cytosolic protein was cloned from JURKAT cells and over-expressed in *E. coli* allowing further characterisation to be carried out (Standaert et al., 1990). These authors showed that FKBP was approximately 12kDa, hence this founder member of the FKBP protein family became known as FKBP 12. Isolation, sequencing and PPIase activity of a *Neurospora crassa* homologue of FKBP 12 was reported at the same time (Tropschug et al., 1990).

1.4

STRUCTURAL STUDIES OF CYCLOPHILIN A AND FKBP 12

Structures of native Cyclophilin A and FKBP12 have been reported, as have many structures of complexes of these proteins with small molecule and protein ligands. The structures of firstly human cyclophilin A followed by FKBP12 will now be considered in greater detail.

1.4.1

Structure of human Cyclophilin A

The structure of native human Cyp A was first determined in 1991 to a resolution of 2.5Å (Kallen et al., 1991; Ke et al., 1991). Subsequently, a number of higher resolution Cyp A type structures from not only human but also other species including *Saccharomyces cerevisiae, Escherichia coli,* and the respectively free-living and parasitic nematodes *Caenorhabditis elegans* and *Brugia malayi* have been reported.

All cyclophilin structures available to date share a very highly conserved fold; an anti-parallel 8-stranded β-barrel capped at either end by two α-helices. Figure 6 illustrates the structure of human Cyp A, residues subsequently shown to be important in the interaction with CsA (Pflugl et al., 1993) are highlighted. A comparison of both native and liganded cyclophilins from various species showed the main structural differences between the proteins were to be found in external loop regions (Taylor et al., 1997).

The structure was compared with that of other β-barrel proteins, for example, retinol-binding protein, β-lactoglobulin and superoxide dismutase (Ke, 1992). Unlike other members of this fold family, the centre of the cyclophilin barrel is filled with hydrophobic residues, precluding ligand binding within the barrel core.

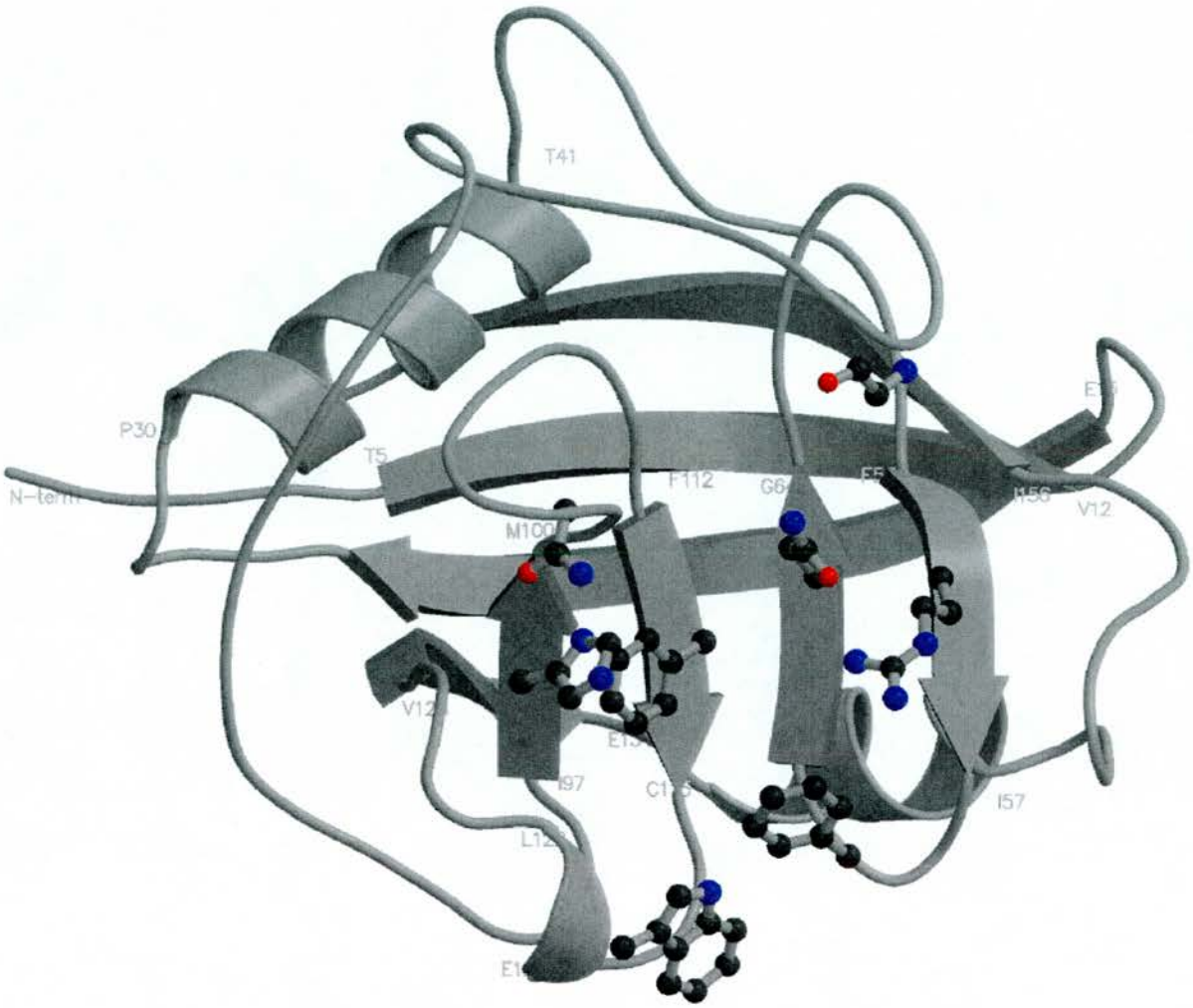


Figure 6

The structure of human cyclophilin A

Helices, loops and β-strands (drawn as arrows) are shown. The residues marking the beginnings and ends of secondary structure elements are labelled. Side chains of some of the residues defining the active site (R55, F60, M61, Q63, G72, A101, N102, Q111, F113, W121, L122, H126) are shown as balls and sticks. The figure was prepared by Molscript (Kraulis, 1991).

A number of publications reporting structures with CsA or small peptide ligands bound to Cyp A suggested the overall integrity of the native structure is maintained following ligand binding. Analysis of structures of Cyp A bound to either CsA or one of eleven CsA analogues revealed a maximum inter structure root mean square deviation (r.m.s.d) of only 0.2Å for the Cα backbone residues 4-67 and 76-160 of the protein (Taylor et al., 1997).

1.4.2
Structure of FKBP 12

The first reported structures of human FKBP 12 were in complex with FK-506 and rapamycin, to resolutions of 1.7 and 2.4Å respectively (Vanduyne et al., 1991a; Vanduyne et al., 1991b). An NMR solution structure of native FKBP was also published in 1991 (Michnick et al., 1991). Once again, numerous structures of bound and free FKBP 12 have been published since the first structures were reported. At the time of writing a search of the PDB structure database found 37 structures of proteins with an FKBP domain alone or ligand bound, however these are mainly FKBP 12.

It is therefore unsurprising that all FKBP structures available to date share a very highly conserved fold, a five stranded anti-parallel β-sheet wrapped with a right handed twist around a short stretch of α-helix, linked together by flexible loops. The structure of human FKBP 12 is shown in Figure 7. Free FKBP 12 has a greater degree of flexibility than bound, particularly within two loop regions, known as the 40's and 80's loops. FK-506 binds in a cavity formed between the β1-sheet and the α-helix, being flanked on either side by the 40's and 80's loops. The structures of bound and free protein, like Cyp A, are also essentially identical as reported in a recent paper (Burkhard et al., 2000).

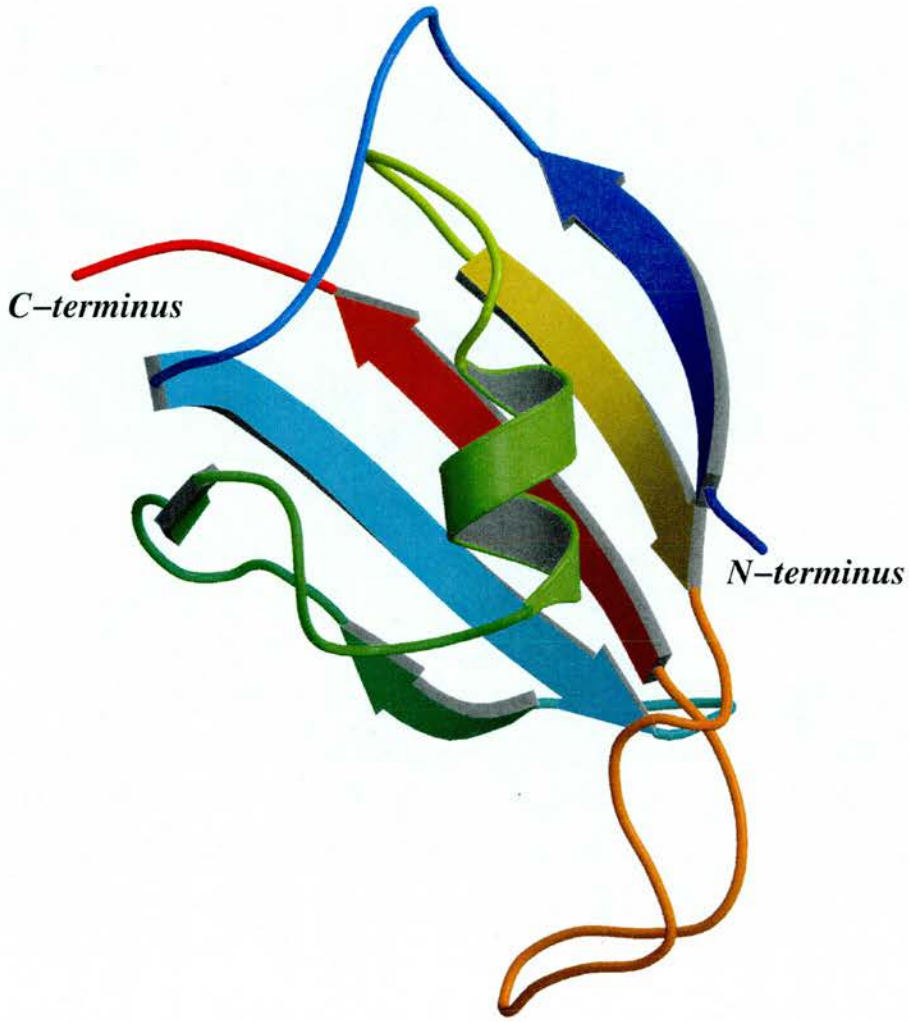


Figure 7

The structure of native human FKBP12. PDB ID no. 1D6O (native FKBP solved to a resolution of 1.85Å) was used as the template molecule. Helices, loops and β-strands (drawn as arrows) are shown. The figure was prepared using Molscript (Kraulis, 1991)

These two defining features of the immunophilin family, high affinity binding to immunosuppressant drugs and possession of PPIase activity were derived from research focussed on two very different areas of biochemistry, illustrating an interesting convergence from diverse fields on this enigmatic family of proteins.

1.5

THE GROWING IMMUNOPHILIN FAMILY

Numerous cyclophilins and FKBP's have been identified since the first reports appeared in the literature. They now constitute an expanding family of proteins, with a very diverse phylogenetic distribution including vertebrates, invertebrates, insects, plants, fungi, bacteria and archebacteria. (for reviews see (Galat and Metcalfe, 1995; Kay, 1996; Gothel and Marahiel, 1999; Ivery, 2000). Immunophilins are also very widely expressed in many different tissues and sub-cellular compartments.

Genome sequencing projects have greatly facilitated the isolation and characterisation of both cyclophilins and FKBP's. Many species contain numerous immunophilin homologues, indeed within a given organism, multiple copies of particular immunophilin types (especially cyclophilins) are often found. Cyclophilin A type proteins remain amongst the most abundant, however both cyclophilin and FKBP domains have been identified as domains of larger proteins involved in very diverse cellular processes.

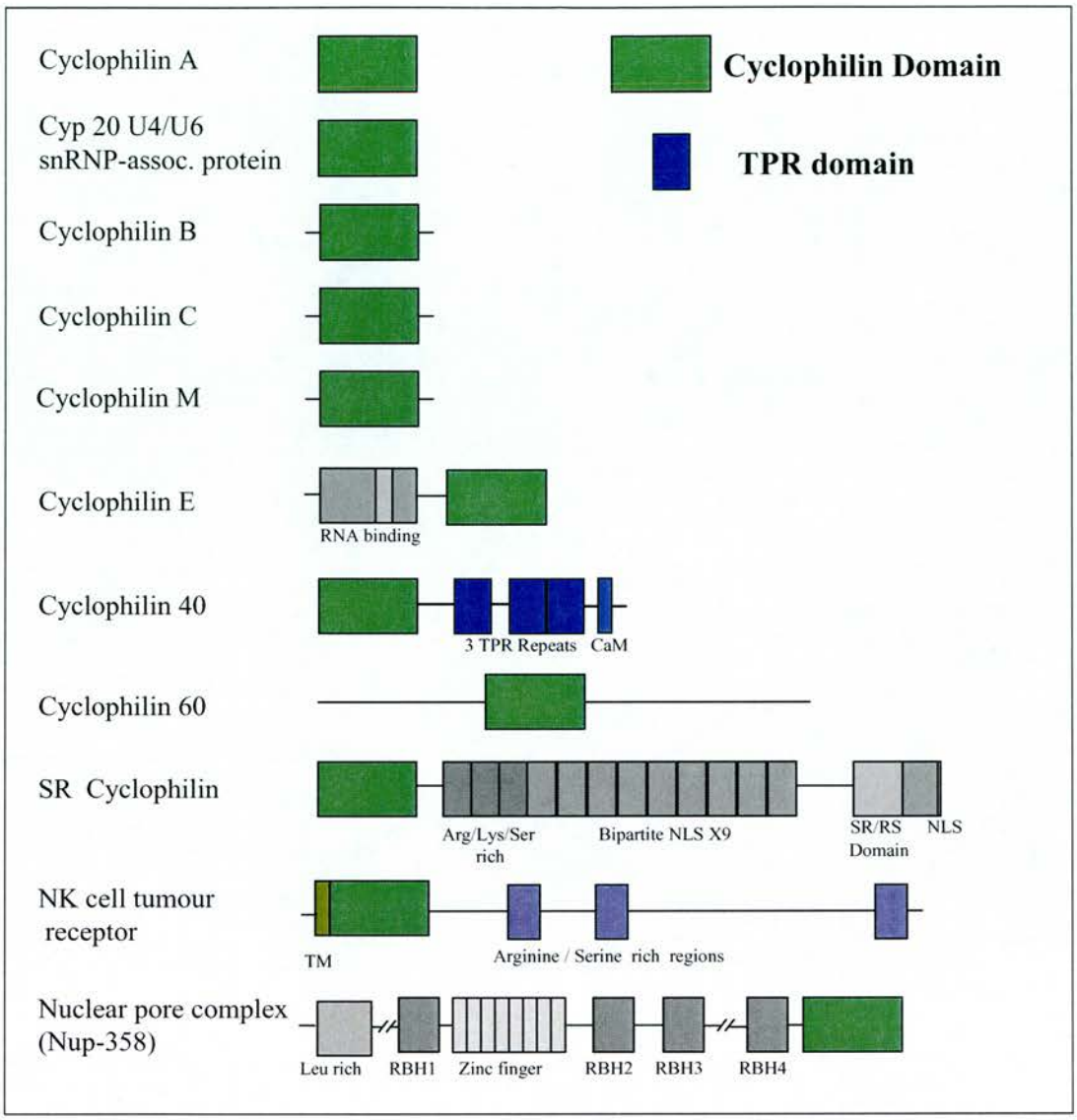


Figure 8

Modular architecture of human cyclophilin proteins.

A search of the SWISS PROT, TREMBL and TREMBL-NEW databases, using "cyclophilin" and "*Homo sapiens*" identified 30 entries, of which 29 were cyclophilins or "cyclophilin like" proteins or fragments. Closer inspection of the sequence, molecular weight and sub-cellular localisation of these proteins allowed broad grouping into the eleven cyclophilin types illustrated above. Green box represents a cyclophilin domain, a blue box indicates a TPR domain. Approximate relative amino acid lengths of each protein are indicated, however the figures are not drawn to scale.

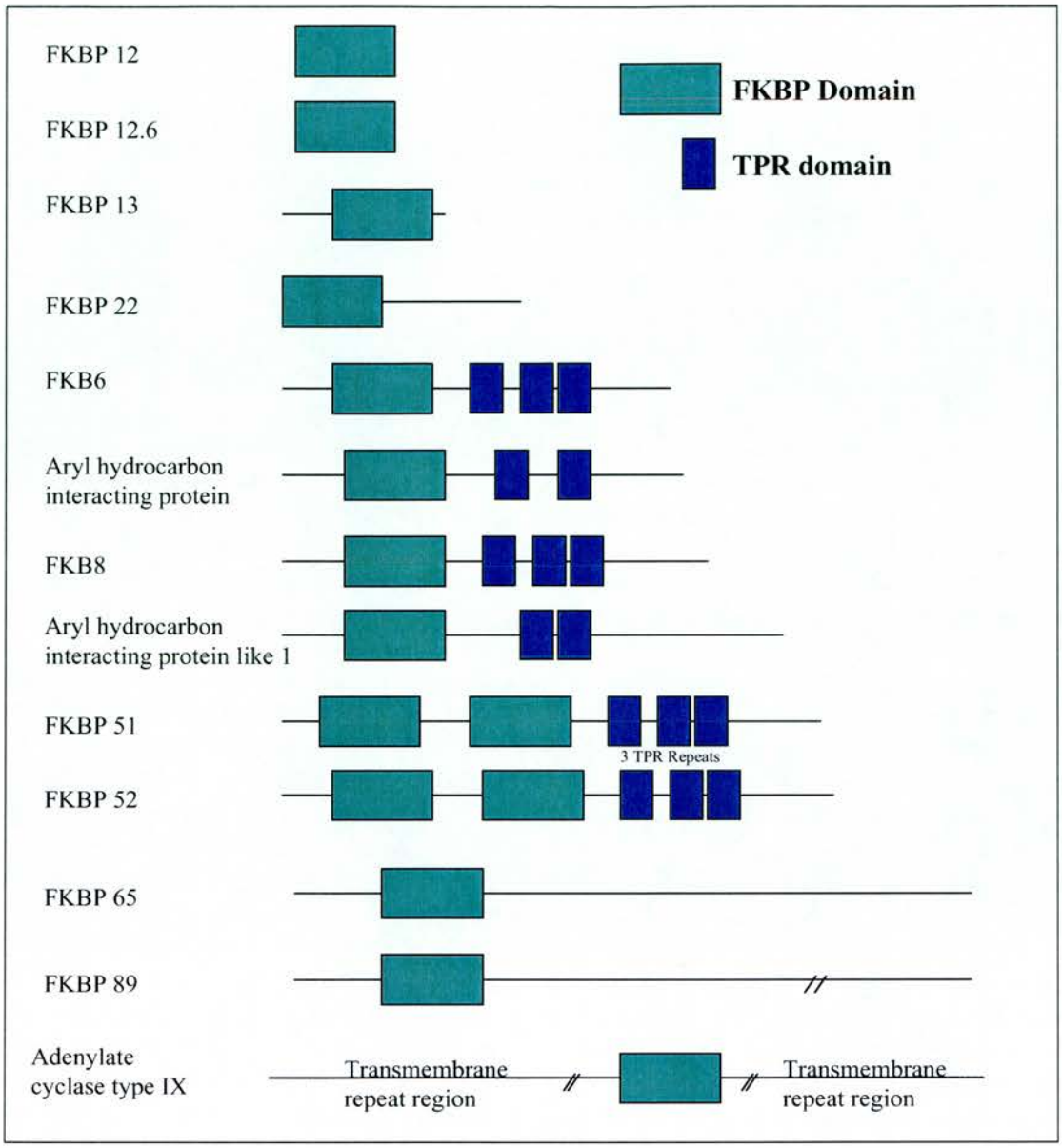


Figure 9

Modular architecture of human FKBP proteins.

A search of the databases with "FKBP" and "*Homo sapiens*" identified 44 entries, of which 30 were FKBPs or "FKBP like" proteins or fragments. Using the same criteria as the previous figure, these proteins could be grouped into the thirteen types illustrated above. Pale green box represents an FKBP domain, a blue box indicates a TPR domain. Approximate relative amino acid lengths of each protein are indicated, however the figures are not drawn to scale.

Isolation of larger FKBP's

A number of reports of larger proteins with affinity for FK-506 appeared in the literature from 1990 onwards, variously called "Hsp56", (Yem et al., 1992), "p59/p60/FKBP59" (Callebaut et al., 1992; Massol et al., 1992; Renoir et al., 1992; Lebeau et al., 1992). These proteins were frequently isolated as components of steroid receptor complexes, often associated with heat shock protein 90, hence they were described as Heat shock protein (Hsp) binding immunophilins. This interaction with Hsp 90 was further studied and was shown to evolutionarily highly conserved, as evidenced by data from experiments using both yeast and mammalian systems (Tai et al., 1992; Tai et al., 1993). Further characterisation of purified FKBP 59 resulted in it being renamed FKBP 52, due to its determined molecular weight (Peattie et al., 1992). The N-terminus of FKBP 52 is comprised largely of three repeats of the FKBP 12 domain. Domain one shares highest amino acid homology with FKBP 12 (55%), domains two and three are increasingly divergent.

The trigger factor of *E. coli* contains an FKBP like domain, both PPIase and chaperone activities have been reported for this protein. The PPIase active site and the high affinity binding site for unfolded protein substrates were shown to be located within different domains of the intact protein, co-operation between the two domains was required for full activity (Scholz et al., 1997).

1.5.1
Immunophilin functions

The functions of these abundant and highly conserved proteins have proved very difficult to determine. In a few cases, however, roles for members of this family have been characterised. A comprehensive although now somewhat out of date review was published in 1995 (Galat and Metcalfe, 1995) which included data on immunophilin function where known.

Some updated examples of immunophilin function will be described in more detail below.

1.5.1.1

Cyclophilins in the pathogenesis of HIV infection

Cyclophilin A and B were isolated in a two-hybrid screen for proteins interacting with the HIV-1 Gag protein (Luban et al., 1993). Cyclophilin A in particular was seen to be a constituent of the HIV-1 virion and was shown to be required for the infectious activity of HIV-1 virions (Franke et al., 1994; Thali et al., 1994; Braaten and Luban, 2001). Numerous studies addressed the possible role of Cyp A in the pathogenesis of HIV-1 infection, however results remained largely difficult to interpret. HIV-1 replication was shown to be inhibited by CsA and its non-immunosuppressive analogues by binding to Cyp A and interfering with its association with the virion (Franke and Luban, 1996; Papageorgiou et al., 1996). A crystal structure of the N-terminal domain of HIV-1 capsid (residues 1–151) in complex with Cyp A was published (Gamble et al., 1996) and revealed a loop on the HIV-1 capsid (formed by residues 85 – 93 in the HIV viral sequence) bound in the Cyp A active site. The *trans* conformation of Pro90 within this loop was unexpected, and was taken to be indicative of the role potentially played by Cyp A in HIV infection. The exact role played by Cyp A in HIV-1 infection is however as yet unclear.

1.5.1.2

Calcium regulated cell signalling

FKBP 12 was first found associated with the calcium release channel/Ryanodine receptor (RyR) of the sarcoplasmic reticulum (Jayaraman et al., 1992) and also with the predominantly endoplasmic reticulum receptor IP_3 which mediates calcium release from the ER (Cameron et al., 1995). FKBP was soon implicated in the regulation of these large calcium release channels, each of which are comprised of 4 identical subunits, one molecule of FKBP binding to each of the four subunits of the tetramer (Timerman et al., 1993). Recent reports have characterised and mapped the FKBP 12 binding site on different isoforms of the RyR and IP_3 receptors (Bultynck

et al., 2001) and also investigated the mode of action of FKBP 12 with this receptor (Gaburjakova et al., 2001)

1.5.1.3
Plant growth and differentiation

The isolation of *pasticcino* mutants of *PAS1* (*pas1, pas2 and pas3*) identified these genes as being involved in control of cell division and differentiation, stages critical to plant development. The gene products were characterised in greater detail (Vittorioso et al., 1998). The product of the *PAS1* gene from *Arabidopsis thaliana* is an FKBP 52 type protein, PAS1 (Faure et al., 1998). The protein was approximately 50% similar to mammalian FKBP 52, shared 4 of the 14 residues involved in the binding to FK-506, with 6 binding residues being those commonly found in other FKBP proteins. The C-terminal region of PAS 1 contains 3 copies of the TPR motif. Mutation in the *PAS1* gene produces plants with a very disordered phenotype and a greatly reduced flowering ability. This study provides an excellent example of a role for FKBP 52 like proteins in plant development, characterisation of partner proteins for PAS1 may provide further insight into their roles.

A role for Cyclophilin 40 in the vegetative but not reproductive maturation of the shoot in *A. thaliana* has been shown (Berardini et al., 2001). An *A. thaliana* Cyp 40 homologue is encoded by the SQUINT gene (SQN). Loss of function mutations in this gene produced plants with a reduced number of juvenile leaves and also subtle changes in inflorescence morphology. The authors suggest these results indicate Cyp 40 plays a specific role in *Arabidopsis* development, regulating the activity of specific signalling pathways.

1.5.1.4
Immunophilins as stress reactive proteins

Recently, a number of papers have been published providing evidence that cyclophilins are involved in responses to and also regulated by, a variety of environmental stress factors, for a short review see (Andreeva et al., 1999).

These factors have included heat, oxidative and heavy metal stress.

Heat shock

A report in which the level of Cyp 40 mRNA was increased 75 fold following exposure of a breast cancer cell line to heat stress was recently published. There was however no corresponding increase in the level of Cyp 40 protein expression, even after prolonged exposure to the elevated temperature, due to an increased rate of turnover of the protein (Mark et al., 2001).

Evidence for FKBP proteins acting in the stress response is more limited, however, the level of wheat FKBP77 mRNA was 14 fold higher under conditions of heat stress (Kurek et al., 1999).

Oxidative stress

Up-regulation of cyclophilin protein production increased by greater than 10% following exposure to oxidative stress. Other "molecular chaperones" were also up-regulated in human umbilical vein endothelial cells (Dreher et al., 1995). In a similar vein, results were presented by another group, investigating the effect of reactive oxygen species (ROS) on vascular smooth muscle cells (Liao et al., 2000). Heat shock protein 90 and cyclophilin B were amongst those secreted in response to oxidative stress. Oxidative stress was reported to regulate cyclophilin function *in vitro* and *in vivo* (Jin et al., 2000). The exact nature of the role of these proteins in ischaemic events remains unclear, however a great deal of research is focussed in this area.

In 1998, Jaschke *et al* published a paper reporting a novel interaction between Cyp A and Aop1 (a member of the family of proteins thought to be involved in defence against the effects of oxidative stress). A yeast two hybrid screen revealed a specific interaction (not substitutable by another cyclophilin). Further experiments suggested Cyp A also stimulated the enzymatic activity of Aop1. This group hypothesised Cyp A acted as a helper protein for Aop1, working together against the damaging effects of oxidative stress (Jaschke et al., 1998). This suggestion is further supported by additional data showing Cyp A is involved in protecting cells against oxidative stress. In this study necrotic cell death was induced by "chemical ischaemia", Cyp A was shown to have a protective role whilst, in this test system Cyp D had a deleterious effect (Doyle et al., 1999).

Up-regulation of cyclophilin mRNA was seen in *Solanum commersonii* (potato) in response to a number of stress events including cold, drought, chemicals and wounding, suggestive of a general role in protecting the plant against various stress factors (MezaZepeda et al., 1998).

Heavy metals

Two cyclophilin genes were isolated from the common earthworm, Cyp A and Cyp B type isoforms. Following exposure to heavy metals, Cyp A mRNA was upregulated 38 fold whilst Cyp B expression remained stable (Sturzenbaum et al., 1999).

1.5.2

Immunophilins as protein folding catalysts and molecular chaperones

Two main classes of "molecular chaperone" activity have been identified.

1. Proteins with enzymatic properties accounting for acceleration of the rate limiting steps in protein folding (protein folding catalysts).
2. That which is functionally responsible for the correct, folding, assembly and transport of newly synthesised proteins in the cell, which also fulfil the accepted definition of a molecular chaperone.

These are frequently overlapping activities and therefore cannot be considered as necessarily mutually exclusive.

1.5.2.1

PPIases as folding catalysts, contributing steric information to the folding process

PPIase activity

The PPIase activity of Cyp A has been shown to be necessary for functional maturation and expression of the α7 homo-oligomeric neuronal nicotinic and type 3 serotonin receptors in *Xenopus* oocytes (Helekar and Patrick, 1997). This was demonstrated by a CsA induced reduction in functional receptor expression and also by mutation of an active site residue (R55A) that also resulted in reduced functional receptor expression. PPIase activity of the cyclophilin from *B. subtilis* has also been

shown to be necessary to allow growth of the organism under starvation conditions (Gothel et al., 1998).

In co-operation with Hsp 70 and Hsp 60, cyclophilins are a component of the mitochondrial protein folding machinery in *Neurospora crassa*, (Rassow et al., 1995). There is evidence that the larger cyclophilin 40 proteins induce conformational changes in the transcriptional factor c-Myb which allows the C-terminal negative regulatory domain of c-Myb to interact with and inhibit its N-terminal DNA binding domain (Leverson and Ness, 1998). Duina *et al*, (Duina et al., 1998a) also suggest a requirement for a Cyp 40 type cyclophilin (Cpr 7) in the negative regulation of the heat shock response in *Saccharomyces cerevisiae*, under both stress and non-stress conditions. A role for the PPIase domain of FKBP 52 in movement of the glucocorticoid receptor (GR) to the nucleus and in an interaction with cytoplasmic dynein (a retrograde motor protein) was proposed in a recent paper (Galigniana et al., 2001).

A number of authors have demonstrated that *in vitro,* cyclophilins accelerate the rate limiting *cis/trans* PPIase steps in the folding pathway of peptide and protein substrates, especially in proline-rich proteins (Lang et al., 1987b; Bachinger, 1987). The importance of such a mechanism to speed up a rate limiting step in protein folding may be particularly important in collagens, due to the abundance of the repeating Gly-X-Pro tripeptide motif (where X can be any amino acid) in these proteins (Schonbrunner et al., 1991; Lang et al., 1987b; Bang and Fischer, 1991; Schonbrunner and Schmid, 1992; Freskgard et al., 1992; Lang et al., 1987a).

FKBP 12 has also been shown to catalyse the folding of protein substrates including RNAaseT1 (Tropschug et al., 1990), carbonic anhydrase (Kern et al., 1994) and FKBP 12 itself (Scholz et al., 1996).

1.5.2.2

Molecular chaperones

Ellis first proposed the concept of molecular chaperones in a landmark paper published in Nature in 1987 (Ellis, 1987). The earliest precise definition of a

"molecular chaperone" was published in 1989 (Ellis and Hemmingsen, 1989). This read as follows "Molecular chaperones are proteins whose role is to mediate the folding of certain other polypeptides and, in some instances, their assembly into oligomeric structures, but which are not components of these final structures".

Proteins belonging to the molecular chaperone family have therefore, amongst other functions, adapted to promote efficient protein folding, interacting with newly synthesised polypeptide chains to minimise inappropriate interactions and maximise the yield of native folded protein. The requirement for assisted folding is thought to arise from the greater incidence of co-translational protein folding found in eukaryotes when compared with post-translational folding that tends to be the norm in prokaryotes. This is thought to be due to the higher incidence of multi-domain proteins in eukaryotes.

Diseases involving protein misfolding/aggregation have also focused considerable research effort upon mechanisms within the cell which operate to minimise or prevent "off-pathway" reactions leading to the adverse effects seen with these diseases.

The initial definition of a molecular chaperone has subsequently been expanded to cover numerous accessory functions associated with assembly of large multiprotein complexes and it is in these roles that larger immunophilins appear largely to function.

"Chaperone activity of immunophilins"

It has been suggested that cyclophilins affect the dynamics of protein-protein interactions *in vitro* by preventing aggregation of early folding intermediates of protein substrates (Freskgard et al., 1992). *In vivo,* several lines of evidence point to a role for cyclophilins in protein folding, transport and assembly. Cyclophilins have also been implicated in protein processing. Substrates of cyclophilin A include collagen (Bachinger, 1987; Steinmann et al., 1991), carbonic anhydrase (Fransson et al., 1992; Kern et al., 1994), calcitonin (Kern et al., 1993), creatine kinase (Ow et al., 2001) and homo-oligomeric alpha 7 neuronal nicotinic and type 3 serotonin receptors (Helekar and Patrick, 1997). Endoplasmic reticulum localised cyclophilins have also

been shown to be involved in processing a number of proteins including transferrin (Lodish and Kong, 1991) and are also important in both prolactin signal transduction and nuclear retrotranslocation (Rycyzyn et al., 2000). The cyclophilin domain from larger proteins has also been implicated in protein processing with various substrates. Two examples include the 150kDa tumour recognition complex in which the N-terminal cyclophilin domain exhibits both PPIase and chaperone activities (Rinfret et al., 1994) and RanBP2 in the visual system, where the cyclophilin domain acts to chaperone red/green opsin. Evidence in support of a role for immunophilins in chaperone functions is mounting, two examples are discussed in greater detail below.

The best characterised example of cyclophilin function *in vivo*, is the 26kDa photoreceptor cyclophilin *ninaA* from *Drosophila melanogaster*. The *ninaA* gene product has been shown to be essential for the proper folding and intra-cellular transport of isoforms of the photopigment rhodopsin, RH1 and RH2, from the rough endoplasmic reticulum through the cytoplasm to the cell surface (Stamnes et al., 1991; Baker et al., 1994; Ondek et al., 1992).

The identification of two new retinal cyclophilin isoforms, types I and II, (Ferreira et al., 1995) provided another example of the role of cyclophilins in the visual system, in this case vertebrate retina cone receptors. The cyclophilin domains were present as part of a larger multi-domain protein, RanBP2. These authors showed that two contiguous domains of this large protein, Ran-binding domain 4 and cyclophilin, acted both *in vitro* and *in vivo* as a specific chaperone for bovine red/green opsin (R/G opsin). The interaction was unlike that seen with NinaA however, the cyclophilin domain did not bind opsin directly but stabilised the interaction between R/G opsin and the RBD4 domain (Ferreira et al., 1996). A follow up paper by the same group (Ferreira et al., 1997) suggested the cyclophilin domain was mediating a modification of the R/G opsins, probably via a prolyl isomerisation step. Proline isomerisation was also suggested as a potential molecular switch, regulated by cyclophilin for the loading of opsin onto RanBP2 during visual pigment processing in cones.

A chaperone function was also suggested for Cyp 40 (Freeman et al., 1996). In this series of experiments Cyp 40 was capable of interacting with an unfolded substrate to maintain it in a intermediate folding competent state, without being able to completely refold the substrate.

1.5.2.3

Immunophilins as highly specific chaperone proteins

Evidence against immunophilins as essential folding catalysts/chaperones was provided in a paper describing the dodecuplet mutant in yeast. This paper reported that deletion of 12 immunophilins (8 cyclophilins and 4 FKBPs) from the budding yeast *S. cerevisiae* (Dolinski et al., 1997) produced very minor phenotypic changes, essentially the sum of the individual mutational effects. This group went on to propose each immunophilin therefore regulates and enhances the functions of a specific number of unique binding partner proteins, as yet largely undetermined, rather than carrying out non-specific chaperone activities.

1.5.2.4

Immunophilins and steroid hormone receptors

A number of studies have implicated members of the larger two domain immunophilin family (Cyp 40, FKBP 52 AND FKBP 51) in the maturation of steroid hormone receptors. These proteins function to maintain the receptor in an activation competent state. They may also be acting as scaffold proteins, bringing together the many components of these multi-protein complexes. Within mature steroid receptor heterocomplexes the large immunophilins compete directly for a common TPR acceptor site in Hsp90 (Owensgrillo et al., 1995; Ratajczak and Carrello, 1996; Nair et al., 1997). A recent report (Reynolds et al., 1999) has linked the incorporation of FKBP 51 into glucocorticoid receptor complexes, at the expense of FKBP 52, to reduced receptor hormone binding affinity, thus providing the first definitive evidence that the TPR-containing immunophilins can alter hormone recognition by steroid receptors. In *Saccharomyces cerevisiae*, Cpr 7 is implicated in the negative

regulation of the heat shock response under both stress and non-stress conditions (Duina et al., 1998a). Cpr6 binds Hsp 90 and shares a 49% sequence similarity with Cpr7 in the C-terminal domain, (Duina et al., 1996a; Warth et al., 1997), however, only Cpr7 is able to functionally support steroid receptor activity in yeast (Duina et al., 1996b; Duina et al., 1998b).

1.5.3
Multi-domain immunophilins

Immunophilins are frequently found as domains of larger proteins. A number of examples have been mentioned earlier in this chapter including Cyp 40, RanBP, FKBP 52 and FKBP 77. Other examples of large proteins containing an immunophilin domain include Cyps 4, 8 and 9 from *C. elegans* (Page et al., 1996) and the 150kDa natural killer cell tumour recognition protein. This protein was isolated as part of a putative tumour recognition complex and contains an N-terminal cyclophilin domain (Anderson et al., 1993). The largest proteins isolated to date (358kDa), which contain cyclophilin domains are members of the Ran binding nuclear pore complex forming proteins, with a C-terminal cyclophilin domain (Yokoyama et al., 1995; Vetter et al., 1999).
Large FKBP containing proteins include type 9 adenylate cyclase that contains one FKBP type domain (Nagase et al., 2000) and numerous trigger factor proteins containing an FKBP type domain.

1.5.3.1
Hsp 90 binding immunophilins

Members of the immunophilin family are frequently found associated with the late phase of steroid hormone receptor maturation (discussed in greater detail in Chapter 5, section 5.3.5). A clear preference for particular immunophilins can be determined for individual steroid receptors, by an as yet undetermined mechanism. The precise nature of the interactions between members of these multiprotein complexes is slowly being dissected out by a variety of approaches. One of the most useful has

been exploitation of bioaffinity interactions between proteins, using deletion/point mutants to determine residues critical for the interactions seen. Structural data on a complex between any of the components of these receptor complexes would provide valuable insight into the precise nature of the interactions.

The immunophilins associated with these receptors share a common domain, containing three copies of the tetratricopeptide repeat motif.

1.5.3.2

The Tetratricopeptide Repeat Motif

The tetratricopeptide repeat is a highly degenerate motif of thirty-four amino acids within which no residue is completely conserved. It was first identified and characterised from the *CDC23* gene product of *S. cerevisiae* and from the product of the *nuc2*$^{+}$ gene from *S. pombe*, as a novel repeating 34 amino acid motif (Sikorski et al., 1990; Hirano et al., 1990). These proteins are both involved in chromosome segregation, controlling regulation of the cell cycle. Around the same time, TPR containing proteins, SSN6 and SKI3, (Smith et al., 1995) were also shown to be involved in RNA synthesis.

The TPR motif is often found in tandem arrays of 3 to 16 units, and has been identified in a wide variety of proteins with disparate functions. This includes mediation of protein-protein interactions, particularly in the assembly of multi-protein complexes. TPR containing proteins include subunits of the anaphase promoting complex, the NADPH oxidase subunit p67^{phox}, transcription factors, the PKR protein kinase and peroxisomal and mitochondrial import proteins (Zachariae and Nasmyth, 1996; Leto, 1996; Ma and Whitlock, 1997; Moir et al., 1997; Tang et al., 1996; Chambraud et al., 1999).

A search of the Pfam Protein Families Database (a large collection of protein multiple sequence alignments), (Bateman et al., 2000) revealed some 605 TPR containing proteins have been isolated and characterised to date. This number can be further broken down into species with 401 proteins in eukaryotes, 25 in Archaea and 179 in bacteria.

1.6

SELECTION OF PROTEIN TARGETS FORMING THE MAIN FOCUS OF THIS THESIS

Requirement for protein structures: Of growing importance is a need to define the structural basis of the interaction between immunophilins and their target proteins, leading to greater understanding of their biological functions. Towards this aim, a number of immunophilin targets have been selected as outlined below, and will be discussed in greater detail in subsequent chapters.

Cyclophilin 3 from Caenorhabditis elegans: In 1996, the cloning and characterisation of eleven cyclophilin genes from the free-living nematode worm *Caenorhabditis elegans* (Page et al., 1996) provided an opportunity to study the various biological roles played by these proteins in this genetically tractable and well characterised organism. Studies on Cyclophilin 3 will be described in Chapter 3 of this thesis.

Cyclophilin 40 and its interaction with Hsp90: Cyclophilin 40 was independently isolated by two groups (Kieffer et al., 1992; Ratajczak et al., 1993). Studies on Cyclophilin 40 will be described in Chapter 4. Preliminary purification and characterisation of regions from the C-terminal domain of Hsp90 will be described in Chapter 5. Chapter 6 will cover modelling studies involving a number of Cyp 40 type proteins and also attempts to model the interaction between the C-terminal domain of Hsp90 with Cyp 40.

FKBP 22: FKBP 22 from *Neurospora crassa* was first isolated from a preparation of *N. crassa* hyphae (Solscheid and Tropschug, 2000). Structural studies on FKBP22 including a partial X-ray structure will be described in Chapter 7.

CHAPTER 2

MATERIALS AND METHODS

2.1
EXPERIMENTAL TECHNIQUES

2.1.1
Protein production and purification

Cell culture media, predominantly Luria-Bertani Medium, (LB) and LB agar were obtained from the in-house media facility (ICMB, University of Edinburgh), ampicillin was also obtained from this facility. Kanamycin and carbenicillin were purchased from Melford Laboratories. Stock solutions of each antibiotic at appropriate concentrations (100mg/ml for ampicillin and carbenecillin, 10mg/ml for kanamycin) were made up in water, filter sterilised, aliquoted and stored frozen at –20°C.

Small scale preps (10ml-250ml) involved growing cells in an appropriately sized container. Large-scale cell culture involved growing cells in two litre flasks, in a shaking incubator. Experimental conditions i.e. temperature and rate of stirring were controlled by setting the shaking incubator (New Brunswick), conditions normally employed were 37°C with agitation at 200rpm.

Occasionally the in-house fermenter facility, run as a service, was utilised in an attempt to produce large amounts of protein from one large (12 litre) prep.

The Gradifrac system from Amersham Pharmacia Biotech (APB) was used for all protein purification protocols described in this thesis. Unless otherwise stated, all chromatography resins and columns used were also sourced from APB.

Electrophoretic analysis was carried out using the MiniProtean II system from Biorad, gels were dried for further analysis and storage using the gel drying system of Promega.

2.2
PREPARATION AND HANDLING OF CLONES

2.2.1
General cloning and expression strategy

The approach taken required firstly; isolation of target gene(s) followed by insertion of the gene into a suitable plasmid vector. Transformation of the plasmid into a suitable *E. coli* strain (chosen for its regulated ability to over-express the target protein at high levels) was the final step.
Cloning of all the proteins used throughout the duration of the work described in this thesis was carried out by collaborating groups, as set out in the text of chapters 3, 4, 5 and 6. Thanks to those groups for the clones.

Many of the standard laboratory techniques used in connection with the handling and preparation of cells for use and storage are described in a very useful and comprehensive text (Sambrook and Russell, 2000).

2.2.1.1
Miniprep method to recover plasmid DNA

The method as outlined in the QIAprep Miniprep Handbook was followed. Briefly, a single colony was picked into 5 mls of LB, supplemented with antibiotic and grown overnight. The cell suspension was centrifuged for 5 minutes to pellet cells, followed by

1. Resuspension of cells in 250µl of resuspension buffer P1 (50mM Tris-HCl, 10mM EDTA, 100µg/ml Rnase, pH 8.0) and transfer to an eppendorf tube.

2. Addition of 250µl of lysis buffer P2 (200mM NaOH, 1% SDS), followed by gentle inversion to mix.
3. Addition of 350µl of neutralising buffer N3 (3.0M potassium acetate, pH 5.5), followed by immediate inversion.
4. Centrifuge for 10 minutes at top speed in a microfuge.
5. Place QIAprep spin column into a 2ml collection tube. Apply supernatant from step 4 to the QIAprep spin column.
6. Centrifuge for 30 - 60 seconds.
7. Wash spin column by adding 0.5ml of buffer PB and repeat centrifugation.
8. Repeat wash as above, using 0.75ml PE buffer.
9. Discard flow through and centrifuge for an additional minute to remove residual wash buffer.
10. Place QIAprep column in a clean 1.5ml centrifuge tube. To elute DNA, add 50µl of EB buffer (10mM Tris-HCl pH 8.5) to the centre of each QIAprep column, let stand for 1 minute and centrifuge for 1 minute.

2.2.1.2

Storage of plasmid

Plasmids recovered by the above protocol were stored frozen (-20°C) in 20µl aliquots.

2.2.1.3

Preparation of calcium competent cells

The following method was used to prepare calcium competent cells for transformation

1. Pick single colonies of the given *E.coli* strain, into 5-10ml of media without antibiotic, grow overnight.
2. Inoculate 100mls of media with 1ml overnight culture.
3. Grow at 37°C until the OD_{600} was 0.6 – 0.9.
4. Aseptically transfer cells to two sterile centrifuge tubes, leave on ice for 10 mins.

5. Pellet cells, 3000g, 4°C, 5 minutes.
6. Remove the supernatant, add 10ml ice cold sterilised 100mM $CaCl_2$ to each tube, resuspend pellet.
7. Repeat centrifugation step, as above.
8. Remove supernatant and gently resuspend cells in a final volume of 2ml 100mM $CaCl_2$. Pooled cells may be stored for a week at 4°C.

2.2.1.4

Transformation of plasmid of interest into competent cells.

1. Transfer 100µl of calcium competent cells to a sterile eppendorf, leave on ice.
2. Add 2 – 5µl DNA/Plasmid, mix very gently and leave on ice for 60 minutes.
3. Heat shock at 42°C for 2 minutes.
4. Remove tube from water bath and place on ice for 5 minutes.
5. Add 1ml Luria Bertani broth and incubate with shaking at 37°C for 20 minutes.
6. Transfer between 50 and 100µl of recovered cell suspension to LB Agar plates, supplemented with an appropriate antibiotic. Spread with a glass spreader.
7. Leave plate with agar on the bottom to allow the liquid to soak into the media.
8. Incubate overnight at 37°C in an inverted position.

2.2.1.5

Storage of transformed cells

Small overnight cultures were grown; 800µl was aliquoted into a cryogenic vial and 200µl sterile glycerol added. The vial was vortexed to mix cells and cryoprotectant, followed by freezing at –80°C.

2.3

PROTEIN PRODUCTION, PURIFICATION AND ANALYSIS METHODS

2.3.1

Protein production

The availability of milligram quantities of pure, soluble protein is one of the critical factors that must be addressed when embarking upon biochemical and structural characterisation of proteins.

Throughout the course of the work described in this thesis, many different proteins were produced using a standard methodology developed. This method was used to try to ensure that starter cultures for large-scale preps (4-15 litres) were over-expressing the protein of interest. The method developed involved firstly testing small scale cultures for over-expression of the protein of interest, followed by selection of the "best expresser" to be used in setting up a fresh overnight culture (100ml to 1 litre) to use as the inoculum for the large scale prep (4 –15 litres).

The method is described below and was followed with only minor modifications for all the proteins discussed in this thesis.

2.3.1.1

Small scale test expression of cultures

Clones were supplied by collaborating groups either on LB Agar plates or in semi-solid maintenance media. To try and ensure uniformity of protein produced (very important when considering crystallisation trials), it was considered preferable to start with a single bacterial colony expressing the protein of interest. The first step therefore was to plate out the bacteria onto LB agar supplemented with appropriate antibiotic, grow overnight at 37°C followed by isolation of single colonies for further experimentation. Plates were stored at 4°C, sealed with parafilm.

Protocol to test for over-expression of proteins of interest.

1. Streak bacteria onto an LB agar plate supplemented with appropriate antibiotic.
2. Incubate plate overnight at 37°C.
3. Identify single colonies, pick a number of these single colonies individually into 10mls of media (LB, 2XYT, Terrific broth), supplemented with antibiotic as appropriate, using a sterile platinum loop.
4. Incubate the 10ml culture overnight at 37°C in a shaking incubator.
5. Inoculate 10ml fresh media (with antibiotic) with overnight culture. Dilution varied from 1:20 to 1:100.
6. Place the overnight culture on ice/leave at room temperature.
7. Grow cells until the OD_{600} between 0.4 and 0.6.
8. Remove 200µl of culture prior to induction, freeze at –20°C.
9. Induce with IPTG to 0.4mM; incubate for a further 3-4 hours.
10. Remove a 200µl post induction sample; defrost pre-induced sample, spin top speed in microfuge for 5 minutes.
11. Remove supernatant, take a 20µl aliquot to test on an SDS PAGE gel, add 5µl of 5X reducing sample buffer.
12. Resuspend pellets in 80µl deionised water plus 20µl 5X reducing sample buffer.
13. Heat at 100°C for 5 minutes, briefly microfuge, load 20µl onto SDS PAGE gel.
14. Run gels, stain/de-stain and check for over-expression of protein of interest.
15. Determine which colony appears to be producing the most protein, with minimum contaminants.
16. Set up large scale second overnight culture using a 1:100 ratio of the original overnight culture with fresh media and antibiotics.
17. Inoculate large volume (4-15 litres) using the second overnight culture, derived from a culture proven to over-express the required protein.

2.3.1.2

Large scale expression of proteins of interest

Cultures obtained by following the above protocol were used as the starter culture for large scale preps, as mentioned in the previous section. Two culture conditions were

used for the work described in this thesis, a shaking incubator system using 2 litre flasks and utilisation of the 15 litre fermenter facility available within the Institute.

2.3.1.3
SDS-PAGE polyacrylamide gel electrophoresis

The method followed was the SDS-PAGE system developed by Laemmli (1970), using the MiniProtean II system from Biorad. To detect protein, gels were routinely stained with Coomassie Brilliant Blue R-250. This is the staining method most frequently employed, capable of detecting of the order of 0.1 – 1μg of protein per band. Molecular weight determination was assisted by the addition to each gel of a lane containing marker proteins of known molecular weight and composition, allowing an estimate to be made of the amount of recombinant protein produced.

2.3.1.4
Purification of protein of interest

The starting point for all experiments was to ensure the availability of pure, soluble protein. Purification protocols were developed *de novo* for each protein discussed in this thesis, the goal, as always, being to produce the maximum amount of purified protein, using the minimum number of steps. A number of standard approaches were routinely taken and are listed below

The first step in extracting intracellular proteins produced in *E. coli* involves disruption of the cell wall, thus releasing proteins into solution. One of the gentlest methods available to produce this effect involves the addition of hen egg white lysozyme to the solubilised cell pellet, followed by a centrifugation step to pellet unwanted cell debris. Additionally, metal chelators such as EDTA are required to further disrupt the structural integrity of the outer lipopolysaccharides. Disruption of the cell wall does however lead to the protein coming into contact with proteases, resulting in the potential for proteolytic attack on the recombinant protein, with consequent degradation of protein during the purification process. Addition of a number of proteolytic inhibitors to buffers used in the early stages of purification,

coupled with a fast (minimum number of steps) protocol, conducted at 4°C with pre-chilled solutions minimises the risk to protein.

As with previous steps, a standard protocol to obtain clarified cell lysate was developed, and is outlined below.

2.3.1.5

Preparation of clarified cell lysates

1. Cell pellets produced from large scale preps were defrosted and weighed.
2. Pellets were solubilised in lysis buffer, the composition of which was protein specific, however the pH was usually around 7.5 – 8.0 to facilitate the action of lysozyme (see individual chapters for detailed protocols including buffers used). Pellets were solubilised to 10% (w/v) unless otherwise stated in later chapters.
3. Lysozyme was added dry to 0.01% (final concentration) and the mixture shaken and incubated on ice for an hour.
4. An equal volume of water at 4°C was added and the solution shaken hard.
5. Magnesium chloride was added to 5mM, followed by DNase and RNase to a final maximum concentration of 0.01% (w/v).
6. The protein solution was incubated for an hour, followed by the addition of EDTA to a final concentration of 10mM.
7. The cell supernatant was centrifuged for an hour at 20,000rpm, the supernatant frozen at –20°C and the pellets discarded after SDS PAGE analysis had established the protein was predominantly present in the supernatant.

(Steps 5 and 6 were optional)

2.3.1.6

Designing an effective purification strategy

A purification protocol for each protein was developed by firstly collating all available information on the protein of interest. Factors which were considered important at this stage included the pI of the protein (calculated from sequence or measured), the amino acid sequence, degree of solubility, protein stability and whether the protein had a bioaffinity tag. This information was used in an attempt to

develop a suitable purification method for each protein. The methods were subject to frequent revision until an apparently optimum protocol was developed.

The two techniques most frequently employed throughout this project were ion exchange and affinity chromatography.

Ion exchange chromatography

Ion exchange chromatography is dependent on a number of parameters including the pH of the buffer system used and the isoelectric point of the proteins to be separated. Selection of suitable conditions for each protein (resin, pH, ionic strength), made ion-exchange a very useful technique throughout the duration of this project.

Affinity chromatography

The principle of affinity chromatography, as the name suggests, exploits an affinity interaction between two molecules (GST/Glutathione) or metal chelates with specific residues (nickel/histidine residues). Clones of a number of proteins of interest were engineered to contain tags of either GST or six histidine residues, facilitating the purification of the proteins bearing the tags. The tagged protein may carry a cleavage site, allowing removal of the affinity tag subsequent to purification of the protein of interest.

2.3.1.7

Concentration of target proteins

Ultrafiltration was used to concentrate protein samples. Dependent upon sample volume, one of two methods was used. Large volumes (>50ml) used the Amicon stirred cell system, applying pressure through a membrane of a defined molecular weight cut-off. This allowed smaller molecules to pass through whilst retaining the protein of interest.

Smaller volumes were concentrated using centrifugation, with devices varying in size from 0.5ml to 50ml, also containing membranes with a defined molecular weight cut-off.

2.3.1.8

Quantitation of protein

The Coomassie Plus Protein Assay Reagent Kit (Pierce) was used for protein quantitation. The reagent is a ready to use modification of the Bradford (Bradford, 1976) coomassie dye binding colorimetric method for total protein quantitation.

A standard curve was constructed using bovine serum albumin. A number of dilutions of the protein samples were made to ensure it could be measured within the limits of the assay. Protein concentrations were determined from the standard curve and multiplied by the dilution factor as appropriate.

2.4

PROTEIN CHARACTERISATION

2.4.1

Dynamic light scattering

Dynamic light scattering was used to assist in the characterisation of purified proteins in solution. A number of studies have shown a strong correlation between monodisperse solutions and success in crystallisation trials. The correlation between the aggregation state of protein samples and the outcome of crystallisation trials was investigated for fifteen proteins (Zulauf and Darcy, 1992).

In an attempt to utilise this potentially useful technique, each "new" protein target was routinely analysed by light scattering. The information obtained from these experiments included some estimate of the purity of the protein sample, its state of aggregation and thermal stability. Consequently some insight into the suitability of the protein for crystallisation trials was obtained from these studies.

2.4.1.1

Theoretical background to selected results from dynamic light scattering

The following results were considered to be of most interest for the proteins analysed and are briefly described below.

- **Polydispersity**

The polydispersity value of a given sample indicates the standard deviation of the spread of particle sizes about the reported average radius. A high polydispersity value therefore indicates a greater spread in size distribution, hence a less uniform population of protein molecules.

An estimate of the degree of sample polydispersity is very useful as a number of reports in the literature had suggested protein solutions with low polydispersity (monodisperse solutions) are more likely to crystallise than those with high

polydispersity (McPherson et al., 1995). A polydispersity value of less than 1 is a good indicator that the protein is essentially monodisperse in solution.

- **Estimated molecular weight**

The molecular weight of the protein is approximated from the hydrodynamic radius calculated from the measured translational diffusion coefficient, using the Stokes Einstein equation. This calculation involves many assumptions and is based upon an empirically derived relationship between the Radius (R_H) and molecular weight of a number of well characterised globular proteins. It is also assumed the protein is present in an aqueous buffer system, however parameters can be altered to accommodate samples of different viscosities.

- **Baseline**

The baseline represents an estimation of the degree of compliance of data with its theoretical size distribution. A baseline of 1.000 is indicative of the actual distribution being fully resolved, using the assumptions inherent in theoretical size distribution.

- **Sum of Squares**

This is a measurement of the closeness of fit between the experimental data and an autocorrelation function generated from the analysis results.

2.4.2

Mass spectrometric analysis of protein targets

Mass spectrometry was used to further characterise protein samples prior to crystallisation trials. Two mass spectrometry methods were used; Electrospray Ionisation (ES) and matrix assisted laser desorption ionisation (MALDI). These two methods differ somewhat in detail, however both result in the gentle ionisation of large bio-molecules including intact proteins and peptides.

In electrospray ionisation mass spectrometry, the sample is introduced into a continuous liquid stream (typically a mixture of organic and aqueous liquids such as 50:50 methanol:water), the mixture is pumped slowly at high voltage through a hypodermic needle, to electrostatically disperse (electrospray) small droplets which rapidly evaporate and impart their charge onto the protein sample. Large ions are

typically multiply charged, resulting in data being presented as a mass-to-charge (*m/z*) spectrum which can be deconvoluted. Deconvolution sums up the signal intensity into a single peak at the molecular weight of the sample.

In matrix assisted laser desorption ionisation spectroscopy, a large excess of matrix material is co-precipitated with the protein sample by pipetting a very small volume of the mixture onto a metal substrate and allowing it to dry. The resulting solid is then irradiated by very short laser pulses, producing gas phase, protonated molecules. The resultant *m/z* spectrum can once again be deconvoluted to assign peaks.

2.5
CRYSTALLISATION TRIALS OF TARGET PROTEINS

Crystallisation Of Nucleic Acids And Proteins, A Practical Approach (Ducruix and Giege, 1999) was frequently consulted during experimental work described herein and also in the preparation of this section of the thesis.

The goal of any crystallographic project is the repeatable production of single, high quality protein crystals that diffract to high resolution. Achievement of this goal requires consideration of a number of important components of the process.

2.5.1
Getting started

The ideal starting goal for an X-ray crystallographic project is the availability of a large amount (typically 10-100mg) of soluble, purified target protein. The need to have a constantly available supply of protein of reproducible characteristics and quality has frequently been re-stated in a number of comprehensive yet accessible texts available on the crystallisation of biological macromolecules, (Ducruix and Giege, 1999; McPherson, 1999). The concept of "crystallographic grade" purity and quantity of protein must be addressed prior to beginning crystallisation trials. The proteins crystallised during the work described in this thesis did not necessarily conform to these ideal conditions. Their expression and purification has been

outlined earlier in this chapter in section 2.3 and will be more fully described, including deviations from the ideal, in Chapters 3 (Cyclophilin 3) and 4 (Cyclophilin 40). However, the various criteria for the availability of highly purified protein were largely met, therefore crystallisation trials were set up.

2.5.1.1

Designing an effective crystallisation strategy.

Prediction of conditions which will *de novo* crystallise a protein is very difficult, if not impossible as the process is a multi-parametric one, due to the complex nature of biological macromolecules and their associations with crystallisation agents. All available information on the target protein should be collated ie pI, solubility, number of free cysteine residues, nature (hydrophobic/philic), any known ligands/partner proteins, monomer/dimer/oligomer, stability vs time, temperature/pH and integrated into preliminary attempts to crystallise the protein.

2.5.1.2

The theory of crystallisation

The exact mechanisms underlying the process of crystallisation of proteins is still not well defined. As mentioned above, the process is a multi-parametric one, involving the three classical steps of nucleation, growth and cessation of growth. A number of the parameters important in crystallisation of proteins are highlighted in Table 2.

To grow protein crystals, molecules have to be brought into a supersaturated, thermodynamically unstable state. Upon return to equilibrium, the molecules may develop into an amorphous or hopefully crystalline phase. Once supersaturation has been reached, formation of the first ordered aggregates can lead to an initial nucleation phase. This is followed by a crystal growth phase producing crystals with varying degrees of usefulness and finally the cessation of growth phase, which may result from a number of processes including simply depletion of the target protein from the system.

PARAMETERS AFFECTING THE CRYSTALLISATION OF PROTEINS
Intrinsic physico-chemical parameters
Supersaturation (concentration of protein and precipitant)
Temperature and pH
Time (rates of equilibration and growth)
Ionic strength and purity of chemicals used in crystallisation
Biochemical and biophysical parameters
Sensitivity of the molecule to physical parameters listed above ie pH, temperature
Binding of ligands e.g. inhibitors, substrates, peptides
Additives e.g. reducing/oxidising agents
Sample ageing
Purity of macromolecule
Macromolecular contaminants in the sample
Sequence micro heterogeneities e.g. fragments, post-translational modifications
Conformational heterogeneity e.g. flexible domains, oligomers and aggregates
Batch effects

Table 2

Table 2 highlights a number of parameters considered during crystallisation trials of the proteins discussed in this thesis. Table adapted from (Ducruix and Giege, 1999)

2.5.1.3

Methods of crystallisation

A number of techniques are commonly employed to produce protein crystals. Throughout the work described within this thesis, the Vapour Diffusion method was followed.

Crystallisation by vapour diffusion

A droplet of target protein was mixed with crystallisation solution, ("mother liquor"), at a ratio of 1:1 (commonly 2µl protein solution + 2µl mother liquor) and equilibrated against a reservoir solution containing mother liquor. As a result of this dilution of the protein with mother liquor, the reservoir has a higher concentration of crystallising solution than that in the droplet/mother liquor mix. Equilibration

proceeds by diffusion of volatile species (water/organic solvents), until the vapour pressure of the droplet is equal to that of the reservoir. This results in the concentration of all droplet components, including the protein, resulting in the droplet being brought into a state of super-saturation.

Experimental set up for crystallisation

Experiments described in this thesis predominantly used "hanging drop" vapour diffusion, occasionally "sitting drop" vapour diffusion was used in attempts to improve crystal size and morphology.

The following protocol was used for all crystallisation experiments described in this thesis, with minor modification as appropriate for either hanging or sitting drops. The later steps in the process were carried out as quickly as possible, to minimise possible denaturation of the protein.

1. Establish composition of crystallisation solutions to be used, calculate quantities required to make 1ml of desired mother liquors, take note of location and nature of each individual crystallisation solution on an alpha numeric grid.
2. Label a 24 well Linbro plate with the date, protein and plate number, ensuring correct orientation of the plate with respect to its alpha numeric labelling, to ensure agreement with the planned conditions.
3. Place small pieces of plasticine into the top corners of the plate.
4. Apply grease around the wells on the plate using soft white paraffin.
5. Place appropriate components of each mother liquor into appropriate wells, as calculated and outlined in step 1.
6. Mix well components either by trituration or shaking.
7. Equilibrate to required temperature (if appropriate).
8. Apply a very small volume of protein solution (typically 1.5-2µl) to a siliconised 22mm round glass coverslip.
9. Apply an equal volume of appropriate mother liquor to each drop.
10. Carefully lift and invert the coverslip and drop onto the well.
11. Seal by pressing gently on the coverslip.

12. Examine drops immediately post set up, and at regular intervals thereafter, note any changes in appearance ie denaturation, precipitation and crystallisation.

Initial crystallisation trial conditions for "new" proteins routinely included an ammonium sulphate (AS) screen, a polyethyleneglycol 4000 (PEG 4000) screen and "Crystal Screen 1", a commercially available kit containing solutions consisting of the 50 most successful crystallisation conditions, as established from searching published literature. The AS and PEG screens involved screening protein at a defined concentration (a minimum of 10mg/ml was routinely used) against grids varying precipitant concentration (AS or PEG) and pH.

2.5.1.4

Co-crystallisation experiments

The method as outlined above was also used in attempts to co-crystallise proteins with a number of peptide and non-peptide ligands. The molecule to be co-crystallised was introduced into the well as a component of the mother liquor.

2.5.2

Crystal soaking

Soaking is a technique commonly employed to introduce a target molecule (ligand) into a pre-formed crystal lattice. The following criterion should be met if the soaked crystal is to prove useful for subsequent X-ray diffraction studies.

1. The ligand must be soluble to an acceptable concentration (10 – 500mM) in a solution that closely mirrors mother liquor.
2. The ligand must be small enough to reach its target site by passing through solvent channels within the crystal lattice; therefore for larger ligands co-crystallisation is more appropriate.
3. The proposed binding site in the protein of the target molecule must be unoccupied; compounds with high affinity binding to the active site were also more likely to be visible in the calculated electron density.

4. The crystal must be able to withstand both transfer into and prolonged incubation in, a solution containing the ligand. Crystal cracking, loss of edge definition followed by the crystal dissolving all resulted in no/poor diffraction.

A number of experiments were conducted to establish soaking conditions for each crystal/ligand pair, and involved trial and error, combined with close and frequent observation by binocular microscope, of the gross morphology of the crystal undergoing soaking.

2.5.3

Preliminary characterisation of protein crystals

An important aspect of the experiment is to select individual crystals most likely to produce data of the required quality. This selection process begins with careful microscopic inspection of crystals in the drop, followed by crystallographic screening, in which individual crystals are exposed to X-rays and the resultant diffraction pattern assessed for usefulness.

To be useful in the X-ray diffraction studies, crystals should ideally fulfil a number of requirements, as set out below.

- They must be protein crystals, ideally single, free from micro-satellites and of a reasonable size (ideally no smaller than 0.1x 0.1 x 0.1mm).
- Have symmetrical, sharply defined faces and edges.
- Be capable of diffracting to the degree of resolution required to answer questions posed by the experiment.
- The degree of mosaic spread in the crystal should ideally not exceed 1°.
- Tolerate exposure to X-rays, allowing collection of a dataset. The advent of cryo-crystallography (mounting crystals in a protective solution followed by flash freezing in liquid nitrogen) has greatly facilitated collection of datasets from single crystals, due to enhanced survival time in the beam.

2.5.3.1

Crystal handling prior to data collection

Data collection and preliminary processing were carried out under the supervision of members of the Structural Biology Group, ICMB.

Crystals were removed from the drop by "fishing" with an appropriately sized cryoloop, transferred into freezing solution (if required) and flash frozen in liquid nitrogen.

Under a stream of liquid nitrogen, the loop assembly was mounted onto a goniometerhead (a device that has two perpendicular arcs that allow rotation, alignment and centring of the crystal). Maintenance of the crystal in the liquid nitrogen stream was essential to avoid icing and/or thawing of the drop

The crystal was centred in the X-ray beam, and a "test" image shot. Routinely, two frames (with the crystal rotated through 90° between frames) were used to initially estimate crystal quality.

The quality of diffraction data produced was assessed, followed by either discarding the crystal or setting up for a data collection. Dependent upon the diffraction pattern produced from the test frames, several parameters can be optimised at this stage. These include changing the distance from crystal to detector, exposure time and oscillation range per frame, all with the aim of extracting the maximum amount of usable data from each frame.

2.5.4

Data collection strategies

2.5.4.1

Determination of the unit cell and space group of a protein crystal (protein crystal characteristics)

Protein crystals are ordered 3-dimensional arrays of molecules held together by non-covalent interactions. The repeating unit forming the crystal is known as the unit cell, the overall array is referred to as the crystal lattice. Irradiation of a crystal produces a

phenomenon referred to as scattering. Scattering of X-rays by the ordered molecules within the crystal results in the emission of X-rays from the crystal, each having a different direction and intensity. Collection and processing of these diffracted X-rays produces a diffraction pattern. The diffraction pattern can reveal a large amount of information about likely crystal quality, including the degree of uniformity of the crystal lattice, completeness and resolution of data available. Initial assessment of the data was necessary to determine the space group of a crystal to allow determination of a suitable data collection strategy.

2.5.4.2

Handling of crystals prior to X ray diffraction studies

Protein crystals are fairly fragile and at all stages were treated gently and manipulated as little as possible in an attempt to minimise damage. All data collections were carried out on frozen crystals, to prolong crystal life in the X-ray beam. Freezing solutions suitable for each crystals were therefore necessary. An ideal freezing solution will maintain crystal integrity while also minimising the production of ice in the droplet (ice is troublesome as it produces rings which can obscure data). A standard approach taken with reasonable success throughout this project was to supplement mother liquor with 20% glycerol. Crystals were transferred from mother liquor into freezing solution using an appropriately sized cryoloop, incubated in freezing solution for approximately 5 seconds and flash frozen in liquid nitrogen.

Professor Malcolm D. Walkinshaw or Dr. Paul Taylor carried out all data processing.

2.6

PEPTIDYL PROLYL CIS-TRANS ISOMERASE (PPIASE) ASSAY

This assay determines the rate of conversion of *cis* to *trans* of a proline containing peptide, working on the principle that chymotrypsin will cleave the peptide only when it is in the *trans* conformation, resulting in the release of chromogenic dye.
PPIase activity was assayed essentially as described by Fischer et al (Fischer et al., 1989b), with the suggested substrate solvent application as described by Kofron *et al.* The peptide substrate Suc-Ala-Ala-Pro-Phe-P-Nitroanalide (Bachem) was dissolved in trifluroethanol (TFE) with 470mM LiCl to give a 100mM stock solution that was diluted to 4mM immediately prior to use.

The experimental setup was as follows; proteins were diluted into 50mM HEPES/86mM NaCl pH 8.0 (PPIase Buffer) to give a 2000nM stock. In a 1ml cuvette, 10μl of protein (20nM final concentration) was added to 860μl PPIase buffer and incubated on ice for an hour. Chymotrypsin (100μl of a 6mg/ml solution) was added to the cuvette followed by transfer of the cuvette to a thermostatted Peltier cell holder by which the reaction temperature was maintained at 5°C. The sample was allowed to equilibrate until a steady baseline was established. The reaction was started by the addition of 30μl (120nM) of peptide substrate followed by rapid mixing. The absorbance change at 400nm due to the release of p-nitroaniline was recorded using a Perkin Elmer Lambda 20 spectrophotometer, with data collected every 0.1 second over a two minute time period. A number of protein concentrations were tested using the above method. Analysis of the data was carried out using a double reciprocal Lineweaver-Burke plot which gives an intercept at $1/V_{max}$ on the y-axis, and $1/[S]=-1/K_M$ on the x-axis. The slope of the line is $K_M/ V_{max.}$

Cyclosporin A inhibition and ligand binding assays were carried out essentially as described above except that the samples were preincubated for a minimum of 60 minutes with varying amounts of CsA (kindly provided by Novartis AG) or other ligand as appropriate.

2.7

SEQUENCE ANALYSIS, ALIGNMENT AND MODELLING STUDIES

The ExPASy Molecular Biology server (http://www.expasy.ch/) was used as a gateway to a number of useful "proteomics" tools. These included access to the SWISS-PROT AND TrEMBL protein databases and also many useful tools and software packages including sequence similarity programmes (BLAST, Fasta3) and multiple sequence alignment packages (CLUSTALW, Multalin). A number of tools were also available to assist in the primary and secondary structure analysis of target proteins including Compute pI/Mw, PredictProtein and PSIpred.

Protein modelling studies made use of SWISS-MODEL and the SwissPdb-Viewer.

CHAPTER 3

PURIFICATION, CHARACTERISATION AND CRYSTALLISATION OF CYCLOPHILIN 3 FROM THE FREE LIVING NEMATODE CAENORHABDITIS ELEGANS

3.1
INTRODUCTION

In 1996, the cloning and characterisation of eleven cyclophilin genes from the free-living nematode worm *Caenorhabditis elegans* provided an opportunity to study the various biological roles played by these proteins in this well characterised and genetically tractable organism (Page et al., 1996). Data on both temporal and spatial expression patterns for several cyclophilin isoforms was subsequently obtained by the same group, shedding more light onto the roles played by individual cyclophilins throughout the lifecycle of *C. elegans,* (Page and Winter, 1998; Dornan et al., 1999a).

Previous studies had identified cyclophilins from diverse parasitic species including the closely related nematode *Brugia malayi* (Page et al., 1995a, Page et al., 1995b), *Plasmodium falciparum*, (Hirtzlin et al., 1995; Reddy and Ingallinera, 1995; Reddy, 1995), *Leishmania major* (Hoerauf et al., 1997), *Trypanosoma brucei brucei* (Daothi et al., 1998). The potent anti-parasitic effects of sub-immunosuppressive levels of CsA on these parasites had been clearly demonstrated, (Chappell and Wastling, 1992) resulting in a great deal of interest in identifying the mechanisms and pathways responsible for these effects.

Cyclophilin homologues from the free-living nematode *C. elegans*

Sequence alignments of the proteins encoded by *cyps* 1 – 18 showed the *C. elegans* proteins to be highly variable, falling broadly into four categories.

- Category 1 included CYPS-2, -3, and –7, most closely resembling Cyp A.
- Category 2 comprised CYP-6 as a Cyp B type secreted isoform.
- Category 3 comprised CYP-1 and -5 Cyp C type proteins with characteristic mitochondrial localisation sequences.
- Category 4 included CYPS –4, -8, -9, -10, –11, -12, -13, -14, -15, -16, -17 and -18 representing the most divergent Cyp D type, possessing one or more residue change in the CsA binding domain.

Isoform and Type	Size kDa	CsA Binding Domain changes	Signal peptide	Insert	pI
CYP-1 (C)	18	Conserved	Yes	Yes	6.24
CYP-2 (D)	18	A103C	No	Yes	8.23
CYP-3 (A)	18	Conserved	No	Yes	8.88
CYP-4 (D)	58	A103K, W121Y	No	No	9.4
CYP-5 (C)	19	Conserved	Yes	No	8.75
CYP-6 (B)	19	Conserved	Yes	No	5.25
CYP-7 (A)	18	Conserved	No	Yes	8.68
CYP-8 (D)	54	A103R, W121H	No	Yes	9.89
CYP-9 (D)	36	A103K, W121H, L122C	No	Yes	8.6
CYP-10 (D)	16	A103N, W121H, H126Y	No	No	5.27
CYP-11 (D)	20	W121F	No	Yes	6.07
CYP-12 (D)	18	W121H	No	No	7.9
CYP-13 (D)	36	A103C	No	No	5.67
CYP-14 (D)	44	R55S, F60Y, M61V, A101V, A103N, W121Y	No	No	5.45
CYP-15 (D)	71	H126N	No	No	6.44
CYP-16 (D)	25	M61I, W121E,	No	No	6.86
CYP-17 (D)	62	R55Q, F60K, Q63M, A101S, N102D, A103T, Q111L, F113Y, W121P, L122S, H126G	No	Yes	8.35
CYP-18 (D)	24	R55K, F60K, Q63L, A101I, N102D, A103K, Q111L, F113N, W121D, L122R, H126A	No	Yes	6.4

Table 3

Characterisation of the eighteen cyclophilin homologues isolated from *C. elegans*. After Table 2 from (Page et al., 1996). CsA binding domain residue numbering scheme is based upon that of human Cyp A.

A study to biochemically and structurally characterise a number of cyclophilin isoforms from the eleven available was proposed. Detailed information obtained from such a study may have had implications in the design of targeted antiparasitic drugs. Cyclophilin 3 was chosen as the initial target for further study.

The goal of this part of the project was to fully characterise the protein biochemically, enzymatically and structurally.

3.2

MATERIAL AND METHODS

3.2.1

Cloning of CYP-3

All cloning was carried out by Dr. Antony P. Page, Wellcome Centre For Molecular Parasitology, The Anderson College, The University of Glasgow, 56 Dumbarton Road, Glasgow, G11 6NU.

The following is provided for completeness. *cyp-3* was cloned from cDNA using a reverse transcriptase polymerase chain reaction (RT-PCR) method as described previously (Page et al., 1996). Briefly, poly-A+ mRNA was obtained using the Micro-Fast Track Kit (InVitrogen), and cDNA was synthesized using a cDNA synthesis kit (Amersham, Little Chalfont, UK). RT-PCR conditions were as follows; 10pmol of each primer (cyp3F/Nde and cyp3R/EcoRI) 5mM $MgCl_2$; 6.7mM 2-ME; 4.4μM EDTA (pH8); 1mM of each dNTP; 113μg/ml non-acetylated BSA), 2μl of mixed stage cDNA and 1μl Taq polymerase (Advanced Biotechnologies), cycled 20X at 94°C for 30 secs, 60°C for 1 min and 72°C for 1 min in 100μl. The primers included engineered restriction sites (*Nde*I and *EcoR*I) which after restriction digestion, allowed directional sub-cloning into similarly digested pET-5a expression vector (Promega). Ligation and transformation protocols of the Promega pET-5 technical manual were followed. Plasmid DNA was isolated (Plasmid Kit, Qiagen) and the insert was sequenced in both directions by thermal cycle sequencing on an Applied Biosystems Automated Sequencer to validate that no mutations had been

generated by the RT-PCR procedure. The plasmid DNA was then transformed into competent BL21 (DE3 pLysS) cells (Promega) for subsequent recombinant protein expression.

3.2.2

pET expression of CYP-3 in E.coli

The standard protocol adopted throughout the duration of this work was followed (see chapter 2, Section 2.3). A number of experiments were conducted to optimise the yield of protein produced. These included varying the following factors during the cell culture phase of the process:

(1) Temperature, (2) OD_{600} at which the cells were induced, (3) Amount of IPTG used for induction, (4) Time post induction before cells harvested, (5) Rate of shaking/degree of aeration (as estimated from volume/flask).

Little difference could be seen when comparing the yields of protein produced, therefore the standard approach of inducing cells at an OD_{600} of 0.6 with 0.4mM IPTG and harvesting 4 hours post induction was followed.

The cell cultures were centrifuged, supernatant discarded and pellets frozen at –20°C.

3.2.3

Purification of recombinant Cyclophilin 3

The cell pellets were defrosted on ice and solubilised in lysis buffer (10% w/v) containing 50mM HEPES, 5mM benzamidine, 5mM EDTA, 5mM β mercaptoethanol, pH 7.5. Lysozyme (ICN/FLOW) was added to 0.1% w/v and incubated on ice for 1 hour. An equal volume of deionised water at 4°C was added and the cell suspension mixed. Magnesium chloride to a final concentration of 5mM was added, followed by the addition of RNase and DNase to a final maximum concentration of 0.01% (w/v). EDTA was added to a final concentration of 10mM. The lysed cells were centrifuged, the resulting supernatant was usually frozen and stored at -20°C. To begin the process of purification, the supernatant was defrosted

slowly on ice and the pH adjusted to 6.8, using 2M HEPES, pH6.5. The cell extract was filtered (glass micro-filter through to 0.45μm filter). Protein extract was purified to homogeneity using two purification steps. SP Sepharose cation exchange chromatography with the following buffers - Buffer A: 50mM HEPES, 5mM EDTA, 5mM β-Me pH 6.8 and Buffer B; Buffer A + 0.5M NaCl (pH 6.8). The protein was bound in batch mode to pre-equilibrated resin, incubated for an hour with frequent mixing and poured into an XK 26/20 column. Column fractions were tested on 15% SDS-PAGE gels and fractions containing the protein of interest were pooled and dialysed, using dialysis tubing with a 3kDa cut-off, overnight against buffer A. Protein solution was filtered through a 0.2μm filter and applied to Resource S resin. CYP-3 was eluted from this resin with the following buffers, Buffer A 20mM HEPES pH 6.8, Buffer B, A+ 0.5M NaCl.

Fractions containing CYP-3 were pooled and concentrated to 20mg/ml as estimated by the Bradford assay (Bradford, 1976).

3.3

CHARACTERISATION OF CYP-3

3.3.1

Dynamic light scattering analysis of the purified protein

Dynamic light scattering experiments on the purified Cyp 3 were carried out using the DynaPro 801 with microsampler attatchment. A number of protein concentrations were tested, the protein was monitored for stability and tendency to aggregate.

3.3.2

Electrospray ionisation mass spectrometry

Purified protein was further characterised by electrospray ionisation mass spectrometry on a Micromass Platform II spectrometer. The sample was first subjected to a reverse-phase chromatography step, retained protein eluted with a

solution of water/acetonitrile/TFA (0.05% v/v) and introduced online into the instrument stream.

3.3.3

Peptidyl Prolyl Isomerase (PPIase) assay

The experimental setup was as described in Chapter 2, Section 2.6. Cyclosporin A inhibition assays were carried out essentially as described above except that the samples were pre-incubated for a minimum of 60 minutes with varying amounts of CsA (kindly provided by Novartis AG).
Ala-Pro ligand binding studies were carried out in the same manner.

3.3.4

Crystallisation of Cyclophilin 3

The hanging drop vapour diffusion method as described in chapter 2, section 2.5.1.3 was used throughout this series of experiments.

3.3.5

Data collection

Crystals were transferred from the hanging drop using an appropriately sized cryo-loop into cryo-protectant solution containing mother liquor plus 20% glycerol. Crystals were removed from the cryo-protectant solution after 5 seconds and mounted in a cryo-loop and flash frozen in liquid nitrogen. To establish if the crystal was of sufficient quality to merit a data collection, 2 frames were shot, with 90° between images. X-ray data from suitable crystals were collected at 100K using a MAR image plate and processed using DENZO (Otwinowski and Minor, 1993).

3.4

STRUCTURE DETERMINATION OF THE TETRAGONAL FORM OF CYCLOPHILIN 3

3.4.1

The structure was solved and refined by members of the Structural Biochemistry Group, ICMB, University of Edinburgh. Dr. Paul Taylor first solved the structure. Refinement and later ligand binding studies were carried out by Dr. Su-ying Wu as part of her PhD programme.

Structure solution and refinement

The structure was solved by molecular replacement using the program AMoRe (Navaza, 1994) with the human cyclophilin A structure as the model. Positional and B-factor refinement were performed using SHELX-97 (Sheldrick, 1997), water molecules were added using SHELXWAT (Sheldrick, 1997).

3.5

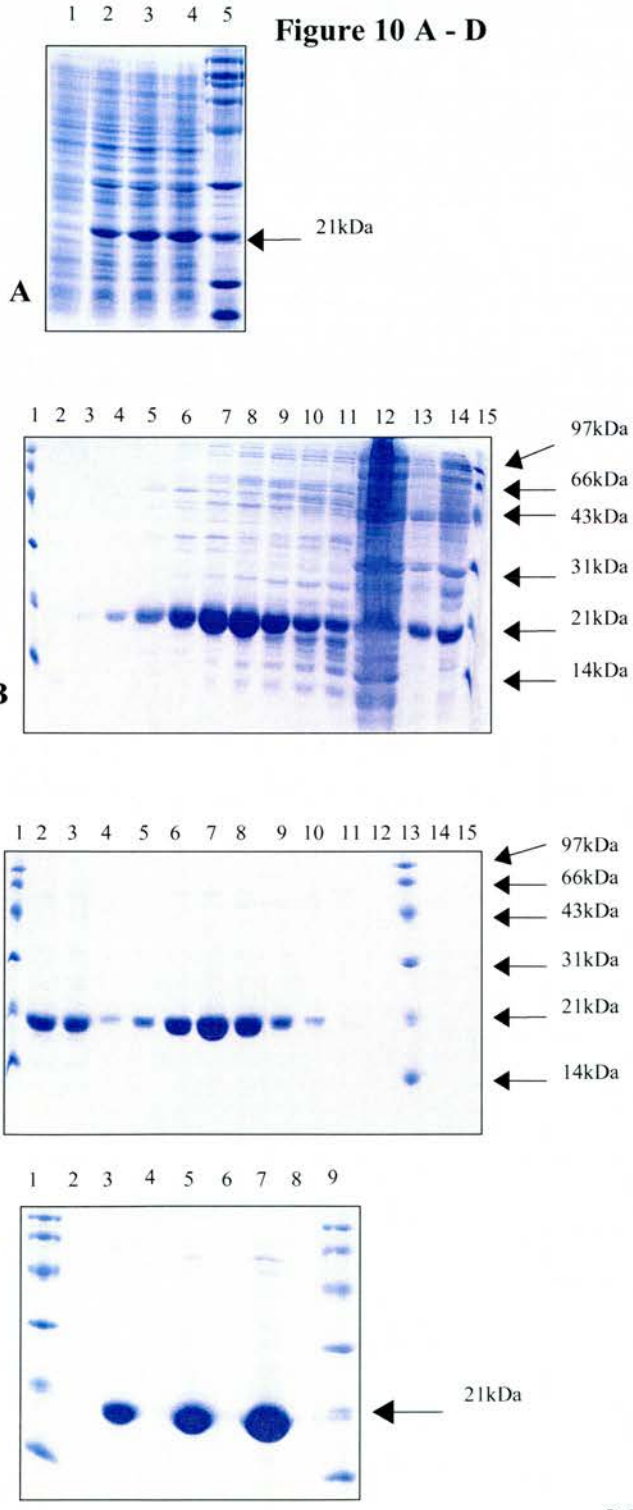
RESULTS AND DISCUSSION

3.5.1

Purification of recombinant Cyclophilin 3

Figure 10 shows SDS PAGE gels illustrating a representative purification run. Conditions were subject to frequent revision, resulting in an optimised purification protocol, utilising 2 cation exchange steps.

Figure 10 A - D



(A)
A representative gel illustrating successful induction of over-expression of CYP-3 by IPTG. Lane 1 uninduced cell culture, lanes 2, 3 and 4, 2, 3 and 4 hours post induction respectively. Lane 5 molecular weight markers.

(B)
Representative gel illustrating a typical elution profile from SP Sepharose. Lanes 1 and 15, molecular weight markers. Lanes 2 – 11, elution of CYP-3 by an NaCl gradient, lane 11 unbound material, lanes 13 and 14 "start material" (clarified cell extract), filtered and unfiltered respectively

(C)
Representative gel illustrating a typical elution profile from Mono S resin. Lanes 1 and 13, molecular weight markers, lanes 2 and 3 material post SP Sepharose and dialysis, unfiltered and filtered respectively. Lanes 4 – 12 elution of CYP-3 by an NaCl gradient.

(D)
Representative gel illustrating purified and concentrated CYP-3. Lanes 1 and 9, molecular weight markers. Lanes 3, 5 and 7 0.5µl, 2µl and 5µl of CYP-3 loaded (lane 7 equivalent to approximately 75µgs of protein).

3.5.2

Dynamic light scattering analysis

Table 4 illustrates a typical data-set obtained from a dynamic light scattering experiment carried out using purified CYP-3. As outlined in Chapter 2, measurements were taken at room temperature using a DynaPro 801 instrument.

Measurement number	**Polydispersity (nm)**	**Estimated Molecular weight**	**Baseline**	**SOS Error**
1	0.5	19kDa	1.003	1.532
2	0.1	18 kDa	1.000	1.223
3	0.1	18 kDa	1.001	1.216
4	0.3	17 kDa	1.001	1.064
5	0.7	17 kDa	1.001	1.354
6	0.5	20 kDa	1.006	1.174
7	0.4	18 kDa	1.000	0.750
8	0.4	18 kDa	1.000	0.753
9	0.5	17 kDa	1.000	1.447
10	0.2	19 kDa	1.004	1.476
11	0.5	17 kDa	1.000	0.897
12	0.1	17 kDa	0.999	0.763
Mean values	**0.4**	**18 kDa**	**1.001**	**1.135**

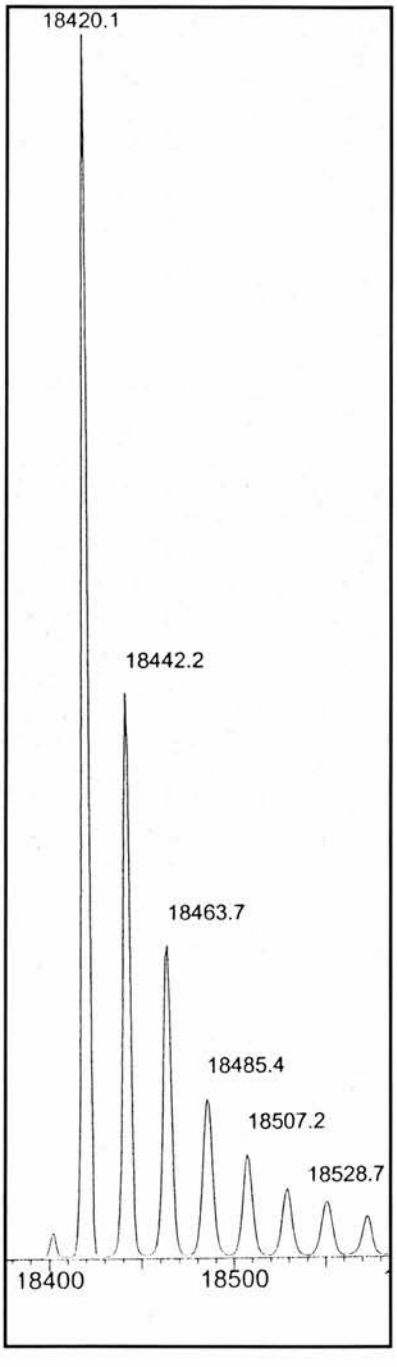
Table 4

Purified CYP-3 at 2 – 10 mg/ml was analysed as described in Chapter 2, using the room temperature micro-sampler attachment of the DynaPro 801.

The four results parameters highlighted above were considered to be most important in the initial analysis of an "unknown" protein sample. As can clearly be seen from the results, the closeness of fit between the experimental and theoretical data was very high. These results therefore suggested the protein was essentially of a uniform composition (mono-disperse) and present in solution as a monomer and was therefore an excellent candidate for crystallisation trials.

3.5.3

Mass spectrometric analysis of CYP-3



The protein was further characterised by electrospray ionisation mass spectrometry on a Micromass Platform II spectrometer. Mass spectrometric analysis was carried out by Dr. Juan Zou as part of a collaborative project investigating the binding of metals to CYP-3. The sample, native protein at 5pmol/μl in a solution of 50% acetonitrile in water with 0.2% formic acid added was introduced into the instrument stream via an infusion pump, the m/z spectrum deconvoluted using MaxEnt software.

Figure 11

Mass spectrometry gives a molecular mass of 18,420 Da. The calculated molecular mass of the sequence is 18,550 Da and suggests that the N-terminal methionine residue is removed post-translationally. The calculated molecular mass of the sequence omitting the N-terminal methionine is 18,419 Da. Additional peaks seen are likely to be sodium adducts of the major protein peak.

3.5.4

Crystallisation of CYP-3

Two crystal forms were grown from the protein isolated using the purification protocol outlined in section 3.2.3. The standard hanging drop vapour diffusion method as described in Chapter 2 Section 2.5.1.3 was followed. Following the establishment of crystallisation conditions, in attempts to increase crystal size, sitting drop vapour diffusion was tried on a number of occasions.

Initial hints as to the "crystallisability" of the protein were obtained from promising, shaped precipitates. Further refinement of these initial conditions, resulted in two very different crystal forms, thin plates and tetragonal bipyramids. The process of refinement of crystallisation conditions involved incremental changes to pH and precipitant concentration, followed by close analysis of the drops for any changes. A number of additives were also tested. The first crystals produced at 4°C were thin plates, conditions were refined at 18°C to produce crystals of greater size and better morphology.

Photographs illustrating sequential changes in crystal morphology are reproduced in Figure 12.

Details of the refinement of crystallisation conditions that produced large single crystals suitable for X-ray diffraction analysis are shown in Table 5.

3.5.5

Preliminary characterisation and data collection

Crystal quality is most important in determining the quality of data collected, however other factors such as exposure time, required oscillation range and crystal to detector distance can also influence the success of a data collection. Parameters such as those above were optimised to produce usable data to the highest resolution possible for a given crystal.

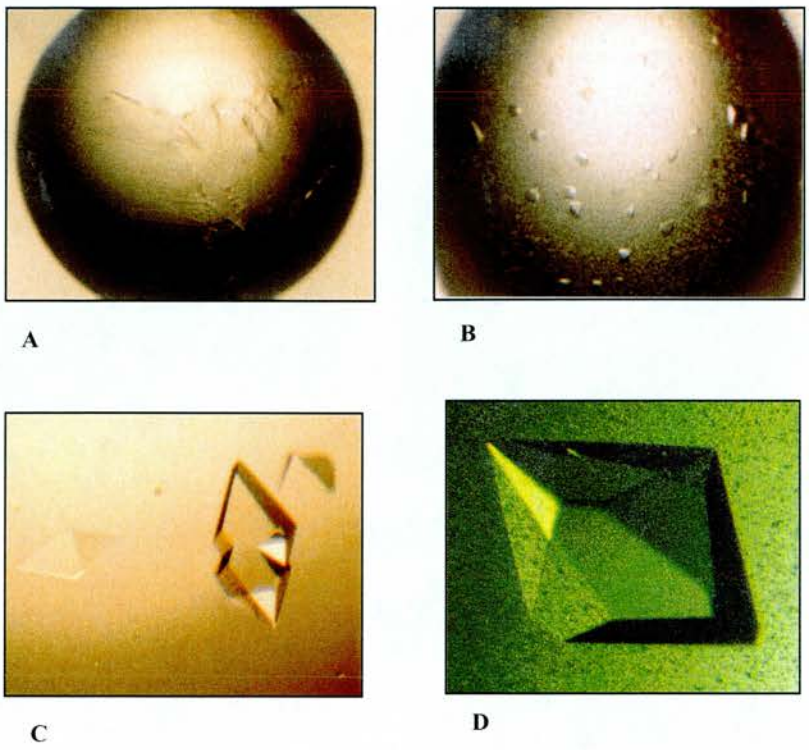


Figure 12

Photographs showing changes obtained in CYP-3 crystal morphology during the process of refinement of crystallisation conditions

Figure 12A

Thin plate form: 4µl hanging drops consisted of 50mM citric acid/sodium citrate pH 6.0, 13% methoxypolyethylene glycol 5000 (MPEG 5000). Protein concentration was 14 mg/ml. The drop was suspended over a well containing 100mM citric acid/sodium citrate pH 6.0 and 26% MPEG 5000, an incubated at 4°C. Thin plate crystals with dimensions 0.5 X 0.2 X 0.005mm grew within a few days. X-ray data were collected from one such thin plate crystal. These crystals were orthorhombic and diffracted to 1.9Å in one direction, with unit cell dimensions of a=58.9 Å, b=67.5Å, c=76.4Å. No further studies were carried out using crystals of this form as they were severely anisotropic.

Figure 12B

Small tetragonal bipyramid form: 4µl hanging drops consisted of 50mM sodium citrate pH 5.6, 15% MPEG 5000 7.5% DMSO. Protein concentration was 7mg/ml. The drop was suspended over the well containing 100mM citric acid/sodium citrate pH 5.6, 15% DMSO and 30% MPEG 5000, at 18°C. Small tetragonal bipyramid

crystals with dimensions 0.1 X 0.05 X 0.05mm grew overnight. X-ray data were also collected from these small crystals. Crystals typically diffracted to 2.5Å, with the same unit cell dimensions as the larger crystals.

Figure 12C

This photograph illustrates two small tetragonal crystals that had formed as a twin. Twinning of crystals presented little problem, as they predominately grew as single crystals. The crystals seen in this photograph are also of intermediate size (0.15 X 0.1 X 0.1mm), useful for data collection at medium to high resolution (maximum resolution on the home source was 2.2Å). Crystal size was a limiting factor in obtaining higher resolution data, therefore sitting drop vapour diffusion was tried in an attempt to obtain larger crystals.

Figure 12D

Large tetragonal bipyramid form: 4µl sitting drops consisted of 50mM sodium citrate pH 5.6, 15.5% w/v (MPEG 5000). Protein concentration was 7mg/ml. The drop was placed onto a micro-bridge into a well containing 100mM sodium citrate and 31% MPEG 5000 and incubated at 18°C. Large tetragonal bypyramid crystals with dimensions up to 0.2 X 0.2 X 0.2mm grew within a few days.

Protein (concentration)	Temp.	Buffer pH	Precipitant	Additives	Crystal morphology	Data collected	Unit cell	Space group	Structure solved
CYP-3, 12mg/ml	4ºC	100mM HEPES pH 7.5	25% PEG 4000	5% DMSO	"Shaped" precipitate	No	No	No	No
CYP-3, 14mg/ml	4ºC	100mM Na Citrate pH 6.0	26% MPEG 5000	None	Thin plates	1.9 Å in one direction	a=58.9Å b=67.5Å c=76.4Å	Unit cell data only	No, crystals severely anisotropic
CYP-3, 20 and 10mg/ml	18ºC	100mM Na Citrate pH 5.6	30% MPEG 5000	5% DMSO	Shower of micro-crystals	No	No	No	No
CYP-3, 7mg/ml (Hanging drop)	18ºC	100mM Na Citrate pH 5.6	30% MPEG 5000	15% DMSO	Small tetragonal bipyramids	2.2 Å	a=b=60.6Å c=123.12 Å	$P4_12_12$	Yes, using Molecular Replacement
CYP-3, 7mg/ml (Sitting drop)	18ºC	100mM Na citrate pH 5.6	31% MPEG 5000	None	Large tetragonal bipyramids	1.8Å	a=b=60.6Å c=123.12 Å	$P4_12_12$	Yes, using Molecular Replacement

Table 5

Refinement of crystallisation conditions for *Caenorhabditis elegans* CYP-3 protein, producing large single crystals suitable for X-ray analysis. Following establishment of initial conditions, crystallisation was very repeatable.

3.6

CYCLOPHILIN-3 3D STRUCTURE

3.6.1

Structure solution and refinement

Crystals belong to space group $P4_12_12$ with unit cell dimensions a=b=60.60 Å c=123.12 Å. The merging R-Factor is 8.3 % and the data is 97% complete to a resolution of 1.8Å. Results of the final refinement are summarised in Table 6.

Summary of refinement statistics	**Native**
Resolution range (Å)	10 – 1.8
$R_{final\ (all\ data)}$ (F>4σ)	20.81% (19.69%)
R_{free} (F>4σ)	28.47% (26.61%)
Root mean square deviations from ideal geometry	
Bond lengths (Å)	0.031
Angle distance (Å)	0.026
Mean B-factor (Å2)	30.37

Table 6

Summary of the final refinement statistics used in the structure determination of CYP-3

The structure was solved by using the molecular replacement programme AMoRe (Navaza, 1994), with Cyclophilin A as the search model. This model gave a single 10 sigma peak in the rotation function and this rotated model gave a 9 sigma peak in the translation function. The structure was refined using SHELX-97 (Sheldrick, 1997) by Su-ying Wu.

All residues are well defined in the electron density map, the exceptions being 2 residues in the N-terminal region (Met1 and Ser2), 2 residues in the loop region (Gly 87 and Glu 88) and Lys 172 in the C-terminal region. Data analysis indicated the structure to be of acceptable quality.

3.6.2

The Overall Architecture of CYP-3

The overall architecture of CYP-3 is similar to that of known X-ray and NMR cyclophilin structures (Braun et al., 1995), (Taylor et al., 1997) with an anti-parallel 8-stranded β-barrel capped at either end by two α-helices. The centre of the barrel is filled with hydrophobic residues from the side chains of Val 6, Phe 8, Ile 20, Met 22, Ile 63, Ile 69, Leu 105, Phe 119, Leu 121 and Ile 165. These hydrophobic residues pack tightly within the core, excluding the possibility of a ligand-binding site within the β-barrel core. Ligand binding has been shown to occur in the active site groove on the surface of the protein see (Taylor et al., 1997) for review. The CYP-3 structure is very highly conserved when compared with the structure of human cyclophilin A, the best characterised member of the family (Kallen et al., 1991; Ke et al., 1991), (Figure 13). The root mean square deviation is 0.55Å for the 620 corresponding C, N, Cα and O atoms from amino acid residues 2 to 44 and 55 to 171 in CYP-3 and residues 2 to 44 and 48 to 164 in Cyp A.

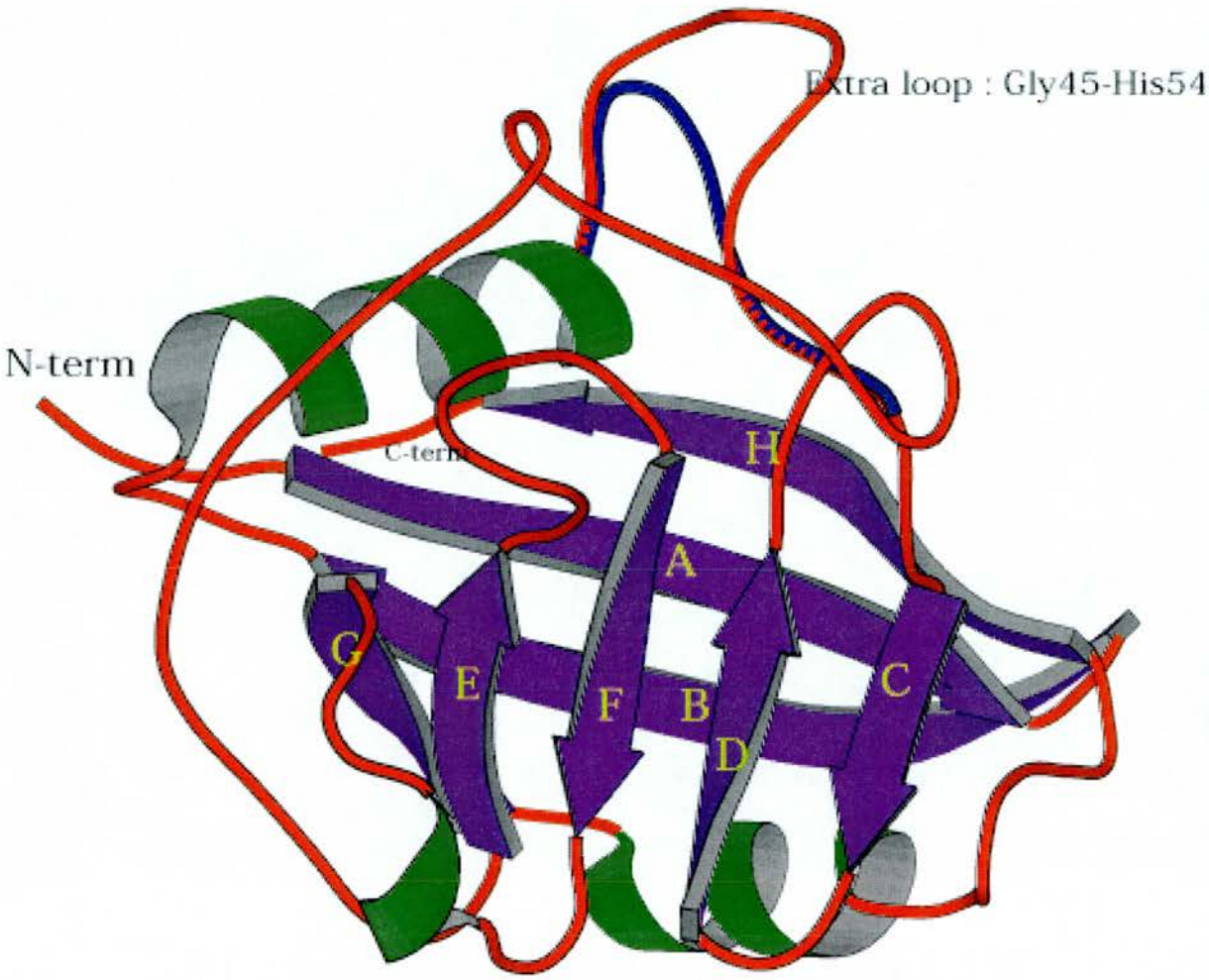


Figure 13

Overlay of the structures of *C. elegans* CYP-3 and human cyclophilin A. The eight labelled strands are shown in blue, the two helices are shown in green, and the additional CYP-3 loop structure is shown in orange. The human Cyp A loop (in blue) contrasts with the shape adopted by the divergent loop, which is formed mainly by 7 additional residues (48KSGKPLH54) in the CYP-3 sequence. The figure was prepared using Molscript (Kraulis, 1991).

3.6.3

Key Structural Features of Cyclophilin 3

Figure 14 shows an alignment of the cyclophilin domains from a number of proteins against human Cyp A. CYP-3 and Cyp A are 67% identical (77% similar) at the amino acid level (Page et al., 1996). The presence of an additional seven amino acid insert in cyclophilin 3 when compared to cyclophilin A, provides the most striking distinguishing feature between these two family members, as can clearly be seen in Figure 13. A number of interesting features have been identified from the structural studies of CYP-3 and these will be discussed in the following section, and are also highlighted in Figure 14, see legend for details

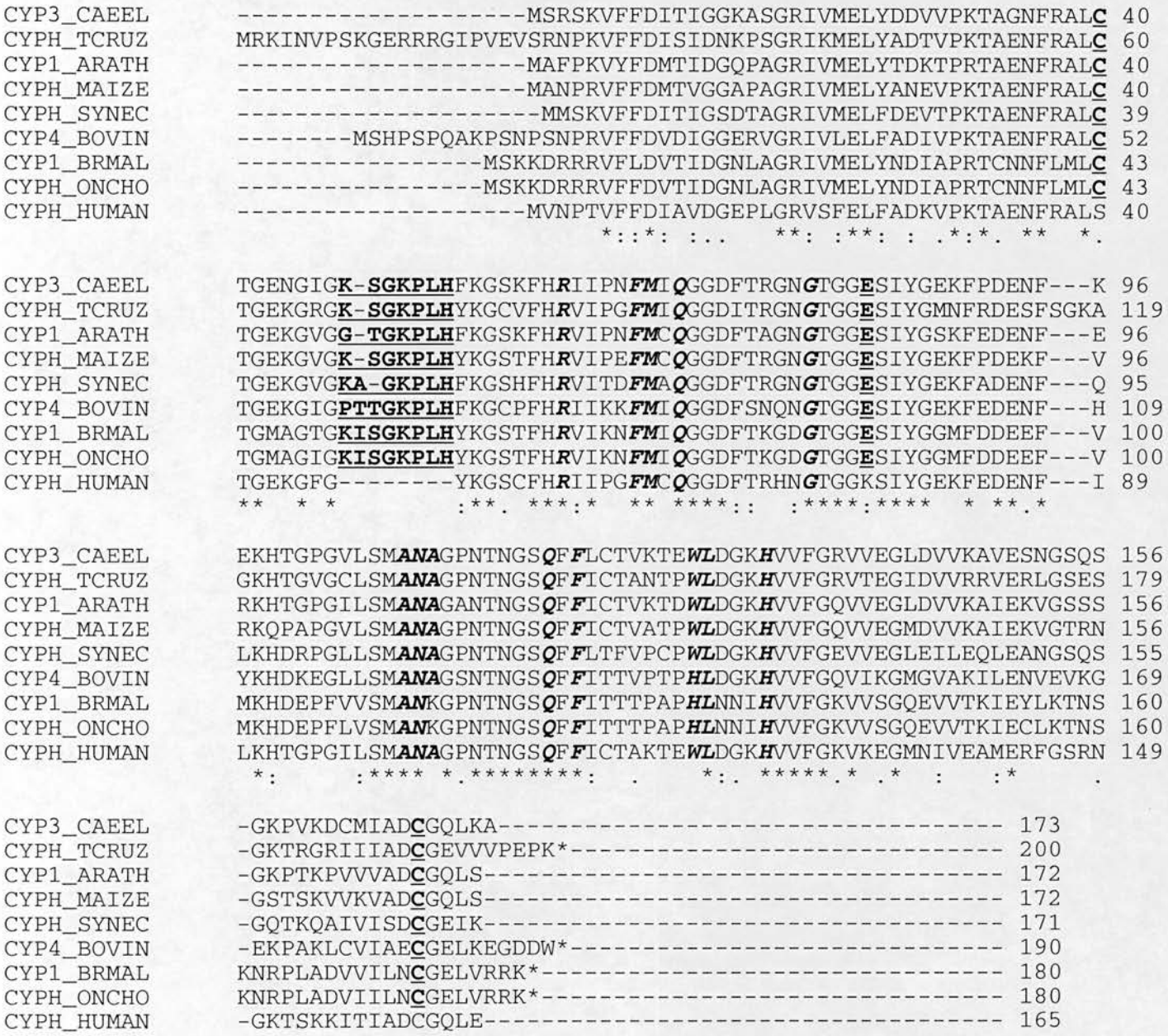


Figure 14

The sequences were aligned by CLUSTALW (Thompson et al., 1994), with dashes introduced for maximum alignment. Inserted residues forming the divergent loop seen in all the aligned sequences with the exception of human Cyp A are in bold and underlined. Cysteine, histidine (His54) and glutamate residues largely but not exclusively conserved within this subgroup of cyclophilins are also shown in underlined bold. Residues important in conferring CsA sensitivity are in bold italics. Symbols below the sequences are part of the results package returned by CLUSTALW. Numbering refers to the amino acid sequence of CYP-3.

3.6.3.1

The active site of CYP-3

The active site of CYP-3 is also very similar to that found in human Cyp A, all 13 residues important in binding to Cyclosporin A being conserved (Pflugl et al., 1994; Page et al., 1996), (Figure 14). The only significant structural difference seen is in the conformation adopted by Arg 62 that places the corresponding side chain atoms 0.7Å apart. The side chains of Met 68 and Lys 125 are also very slightly shifted (Figure 15). This fits with the similar enzymatic activity and inhibition by CsA as is found for human Cyp A. The very close structural similarity of the active site region also suggests that the possible natural cellular substrates are also similar between *C. elegans* and mammals.

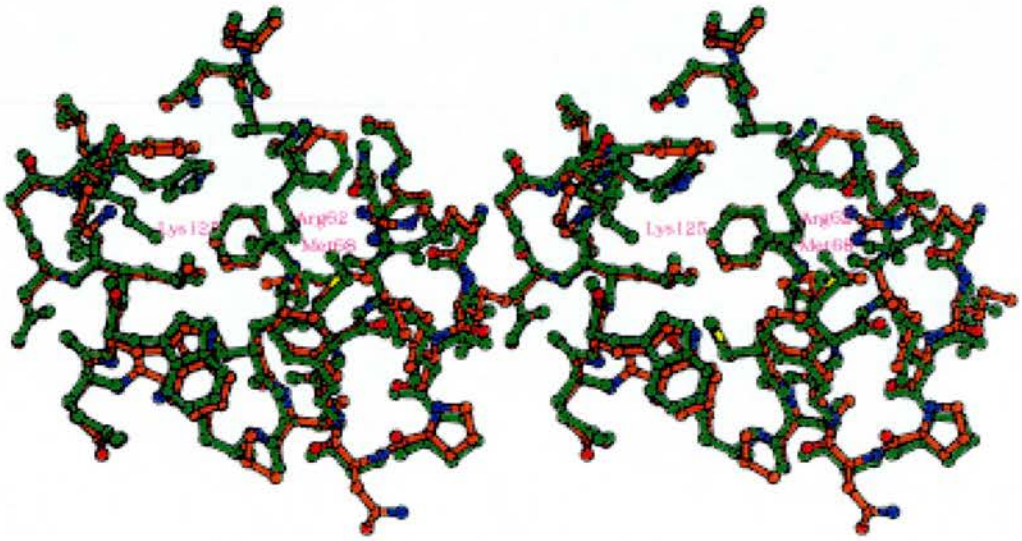


Figure 15

Stereo diagram showing the active sites of CYP-3 (red) and human Cyclophilin A (green). This clearly shows the complete conservation of all 13 residues important in the binding of CsA. The only significant structural difference between these two proteins being in the conformation adopted by Arg 62 that places the side chain atoms 0.7Å apart.

3.6.3.2

The Divergent loop

The most distinctive structural feature distinguishing CypA from CYP-3 is an additional insert of seven residues ^{48}KSGKPLH^{54}, which forms a protruding loop located above the active site of CYP-3 (Fig 13). This feature has also been observed in the Bm-CYP-1 cyclophilin structure from the filarial nematode *Brugia malayi*, (Taylor et al., 1998; Mikol et al., 1998).

The loop is held in place by a very specific network of hydrogen bonds from the side chain of Glu-83. The carboxylic group of Glu 83 forms two strong hydrogen bonds with amide main chain nitrogen atoms of Lys 48 and Ser 49 (Fig 16), resulting in an overall loop conformation which is similar to the most common type 1β-turn geometry.

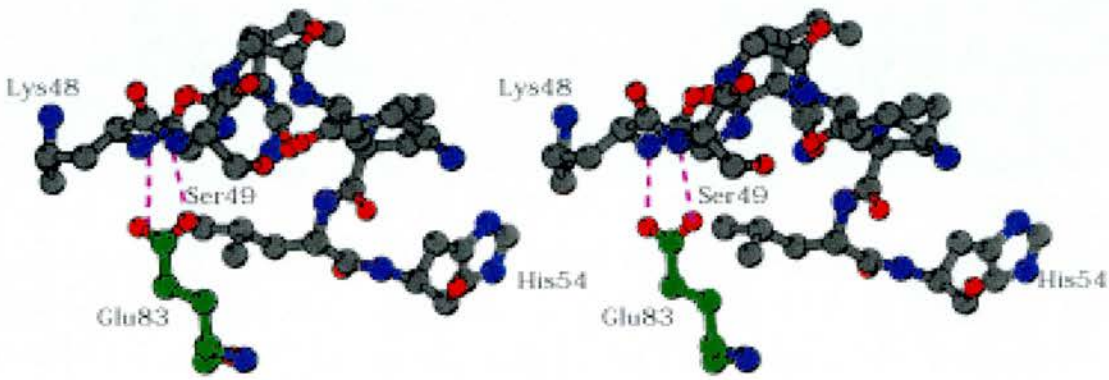


Figure 16

The divergent loop of *C. elegans* CYP-3

The two strong hydrogen bonds from the carboxylic group of Glu 83 to the amide main chain nitrogen atoms of Lys 48 and Ser 49 are shown in violet. The "divergent" loop structure provides a distinctive recognition feature of this subfamily of cyclophilins.

Many cyclophilin homologues with additional loops (including human/bovine cyclophilin 40 and some parasite and plant cyclophilins) have a conserved glutamate in a homologous position. Thus the role of Glu 83 in these "divergent loop" structures is therefore to both lock and tether the loop into a particular conformation. The loop provides a distinctive recognition feature of this sub-family of cyclophilins and may prove to have functional significance in a number of important interactions including protein-protein recognition mechanisms.

Loop structures do however show some divergence as can be seen on comparing the structures from CYP-3 and *B. malayi* (Taylor et al., 1998). The presence of inserts of 6-8 residues has been described previously in a number of members of this protein family, including many plant species (Chou and Gasser, 1997). It is a particularly common feature of the nematode cyclophilins, being found in 7 of the 11 initially characterised from *C. elegans* (Page et al., 1996). A consensus sequence for this additional loop based on multiple sequence alignments is (**GK*LH).

It is interesting to note that this loop also contains part of a conserved ATP/GTP-binding site motif A (A/G-X4-G-K-S/T) (Walker et al., 1982). The high degree of conservation of this inserted sequence over diverse species suggests the loop structure and its locking mechanism are an important functional part of the protein whose role remains to be elucidated, rather than a neutral addition.

3.6.3.3

Conserved cysteines

Conserved cysteines: CYP-3 shares with many other cyclophilins including most divergent loop cyclophilins, such as Bm-CYP-1, two highly conserved cysteines (Figure 14). The sulphur atoms of Cys40 and Cys168 are well defined in the electron density and are separated by 5.38 Å, and are present in the structure as reduced cysteine residues (Figure 17).

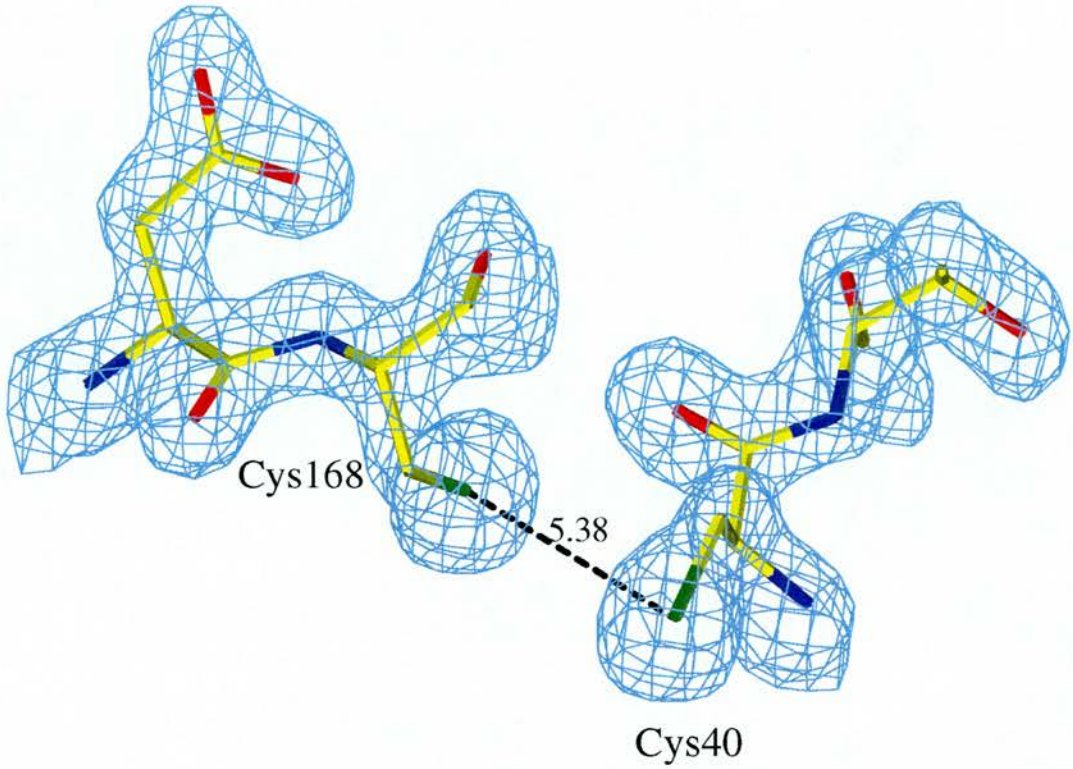


Figure 17

The sulphur atoms of Cys40 and Cys168 are present in the structure as reduced cysteine residues. Torsion angles of the 2 residues are N-Cα- Cβ-S = 53° for Cys168 and -72° for Cys40. The sulphur atom of Cys168 is 3.74Å away from nitrogen Ne2 of His 54, 3.2Å from the carbonyl oxygen of Cys40 and 3.5Å away from the backbone NH of Lys 56. Molecular modelling has shown that a simple rotation about the Cα-Cβ cysteine side chain bonds enables a disulphide bond to form (N-Cα- Cβ-S = -167° for Cys168 and N-Cα- Cβ-S = -134° for Cys40).

These changes from the original staggered conformation to an alternatively favoured staggered conformation results in an S-S distance of 2.1Å, with no other short non-bonded inter-atomic contacts. The role of these conserved residues is unknown. The unexpected presence of the reduced form of these cysteine residues suggests the intriguing possibility that this feature may provide a signalling mechanism in response to oxidative stress. The formation of an S-S bond in an oxidising environment may provide a dynamic signal inducing conformational change(s). Alternatively, the sulfhydryl group of the reduced Cys168 and Cys40 along with the imidazole side chains of His 54 may provide a metal ion co-ordination site.

Interestingly, His-54 (or equivalent) is conserved in many of these divergent loop cyclophilins (Figure 14) and may also be an important part of this putative redox signalling mechanism.

There may be a general link between cyclophilins and a cellular stress response as suggested by the association of human Cyp A with the thiol-specific antioxidant protein Aop1 (Jaschke et al., 1998). This activity is stimulated by the interaction between Aop1 with Cyp A and is not inhibited by CsA. One mechanism for the control of oxidative stress in bacteria is by the protein OxyR which is activated by the formation of an intra-molecular disulphide bond (Zheng et al., 1998). A similar mechanism was proposed for the activation of the Hsp33 chaperone protein which is also redox regulated (Jakob et al., 1999). Oxidation of Hsp33 by hydrogen peroxide causes disulphide bond formation, release of bound zinc and activation of the chaperone activity of Hsp33.

The X-ray structures of CYP-3 and CYP-1 from *Brugia malayi*, (Taylor et al., 1998; Mikol et al., 1998) another member of the divergent-loop subclass of cyclophilins, both show the conserved cysteine residues in the reduced form. Furthermore, His54 which is incorporated in the divergent loop is located 3.6 Å from the sulphydryl group of Cys168. The arrangement of two cysteines and a histidine provides a suitable co-ordination site for zinc. A very similar co-ordination geometry involving two cysteines and a histidine is present in the X-ray structure of a zinc complex of alcohol dehydrogenase (Cho et al., 1997). The arrangement of cysteines and histidines, often present in this newly defined divergent-loop subclass, are therefore potentially involved in either metal co-ordination or in a redox signalling process triggered by the formation of a Cys40 to Cys168 disulphide bond.

3.6.3.4

Temporal and spatial expression pattern of CYP-3

CYP-3 represents one of the most abundantly expressed cyclophilins in the nematode *Caenorhabditis elegans*. This observation is based both on the analysis of relative transcript abundance and due to the fact that it has been isolated 18 independent times during the *C. elegans* expression sequence tag (EST) project ([Anon], 1998). The semi-quantitative RT-PCR analysis of the post-embryonic temporal expression pattern of *cyp-3* indicates that this transcript is most abundant during early larval development, peaking at the second larval stage and dropping off in later development.

Temporal expression pattern

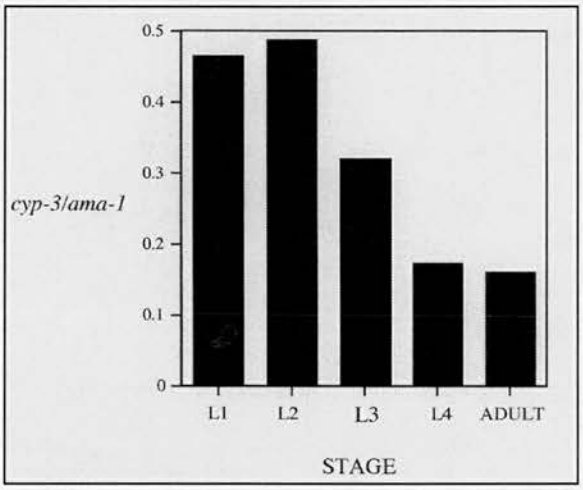


Figure 18

Graphical representation of the semi-quantitative RT-PCR result, illustrating the peak seen for the *cyp-3* transcript in the second larval stage. The temporal abundance of *cyp-3* is expressed as a ration to the constitutively expressed gene *ama-1*. Thanks to Tony Page for the figure.

The expression pattern seen for the *cyp*-3 transcript may correlate with a peak in the amount of newly synthesised structural proteins produced by the developing larvae.

Spatial expression pattern

The spatial expression pattern of CYP-3 was examined by constructing a translational fusion of the promoter region with a reporter gene and staining the transgenic nematodes for β-galactosidase activity. This analysis revealed a very specific tissue localisation for this highly conserved cyclophilin, being expressed exclusively in larval stages in the single celled anterior excretory system. This

reporter construct is tagged to a nuclear localisation signal, hence the specific signal is observed in largest mononucleate cell of *C. elegans*, the excretory cell (White, 1988).

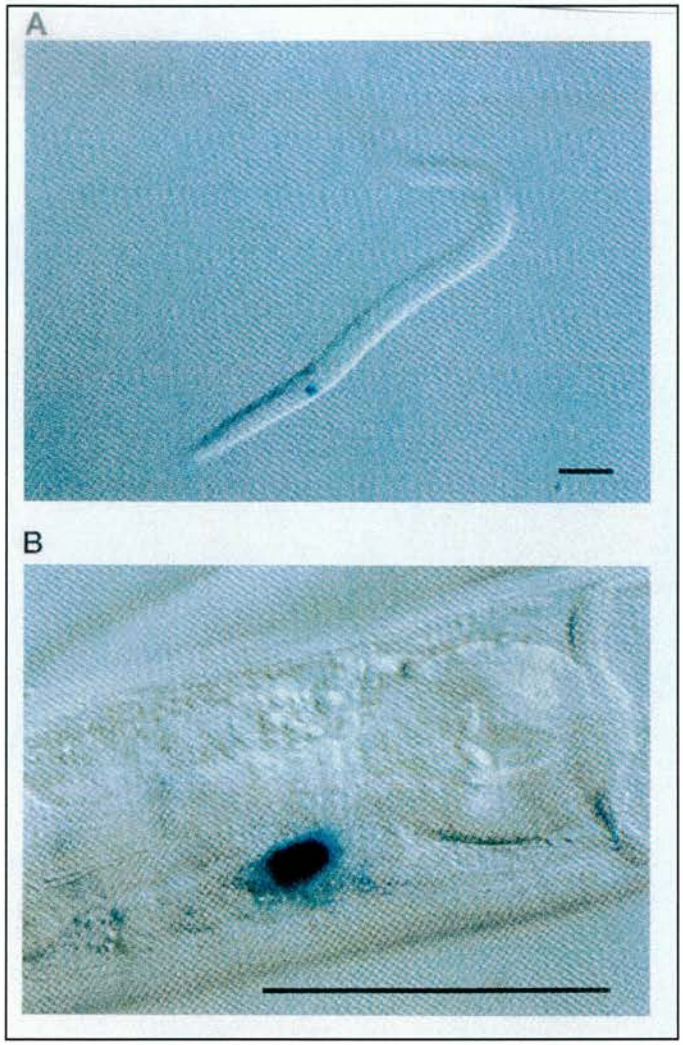


Figure 19 A and B

Micrograph showing the exclusive *cyp-3* reporter gene expression in *C. elegans*. Nuclear localised *β*-galactosidase activity is restricted to the single anterior mononuclear cell. A and B represent individual larval nematodes.

Bar represents 10μm.

Thanks to Tony Page for the figure.

The exclusive excretory cell expression pattern is also intriguing. The excretory system in *C. elegans* is involved in osmoregulation and secretion (White, 1988). The role in osmoregulation is supported by gland cell laser ablation experiments, which resulted in worms becoming bloated and dying in hypotonic medium (Nelson and Riddle, 1984). The importance of the excretory system in secretion is well established for parasitic nematodes, where it is the source of surface coat antigens for parasites such as *Toxocara canis* (Page et al., 1992a; Page et al., 1992b). A role for CYP-3 in either osmoregulation or secretion however, remains to be established.

CHAPTER 3B.

FURTHER STRUCTURAL AND ENZYMATIC STUDIES OF CYCLOPHILIN 3

3.7

EXPERIMENTS FOCUSSING ON CONSERVED CYSTEINE RESIDUES

3.7.1

Metal catalysed oxidation of cysteine residues present in the structure of CYP-3 in the reduced form.

In an attempt to induce disulphide bond formation between Cys40 and Cys168, a number of attempts were made to utilise the ability of metals to catalyse an oxidation reaction. Crystals were soaked overnight in solutions of mother liquor containing a number of different concentrations of either silver nitrate (10mM) or iron III chloride hexahydrate (107mM). The soaked crystals were then flash frozen and exposed to X-rays. Following oxidisation trials, difference electron density maps were calculated, using the native structure as the reference. The difference maps were examined to establish if there had been any changes in the density around residues Cys 40 and Cys 168. No changes could be seen in electron density, therefore no further experiments were carried out.

Attempts using mercury derivatives were carried out as part of a Summer Studentship, held by James Brown. This series of experiments resulted in the crystals either diffracting very poorly or not at all, suggestive of metal perturbing the crystal lattice. Crystallographic aspects of the work described above were carried out by Professor Malcolm D. Walkinshaw, Drs Su-ying Wu and Paul Taylor.

Interest in CYP-3 metal interactions, focusing particularly on the four free Cys thiols, continued within the laboratory.

3.7.2

Anti-arthritic gold complex binding to CYP-3

A number of anti-arthritic gold (I) drugs are in clinical use, their critical target sites are thought to be thiolate sulphurs (cysteine residues), there is however little structural data on adducts of anti-arthritic Au (I) complexes with proteins. Cyclophilin 3 was identified as an excellent template for these studies for two main reasons; the availability of crystals amenable to soaking experiments that subsequently still diffracted well and also the presence of four free Cys thiols identified as likely target sites for metal and in particular Au binding.

Studies using this system with an Au (I) phosphine complex ($AuPEt_3Cl$) were carried out in collaboration with Dr. Juan Zou and others from the group of Professor Sadler, Department of Chemistry, University of Edinburgh. Crystals were soaked with gold complex for time periods that varied between approximately 30 minutes and two weeks. Crystals were significantly altered following soaking with gold complex solutions; cracks appeared and the crystals lost edge definition. Crystals soaked for approximately 30 minutes showed no additional electron density upon calculation of difference maps, whilst those soaked for more than 6 hours failed to diffract at all. Data was obtained from crystals soaked for 2 and 4 hours.

Protein in solution was also reacted with varying amounts of complex, to provide samples for mass spectrometric analysis. Experiments to confirm the accessibility of the four free Cys thiols were carried out. Protein was reacted with the disulphide 5,5'-dithio-bis 2-nitrobenzoic acid (DTNB), changes in the absorption at 412nm were monitored. Native and reacted proteins were also analysed by LC-ESI-MS.

The structure was solved using molecular replacement, with the native CYP-3 structure as the template. Difference electron density maps of crystals soaked for 2 or 4 hours in a 3.5mM solution of gold complex clearly showed a large peak corresponding to a single gold site in the region of an active site residue, His 133. Further rounds of refinement revealed the gold had indeed bound to the Nε2 nitrogen

of His 133 and to PEt_3. Complexation of Au to the active site histidine was also shown to be inhibitory to PPIase activity, as assayed by the PPIase activity assay described earlier in Chapter 2, Section 2.6.

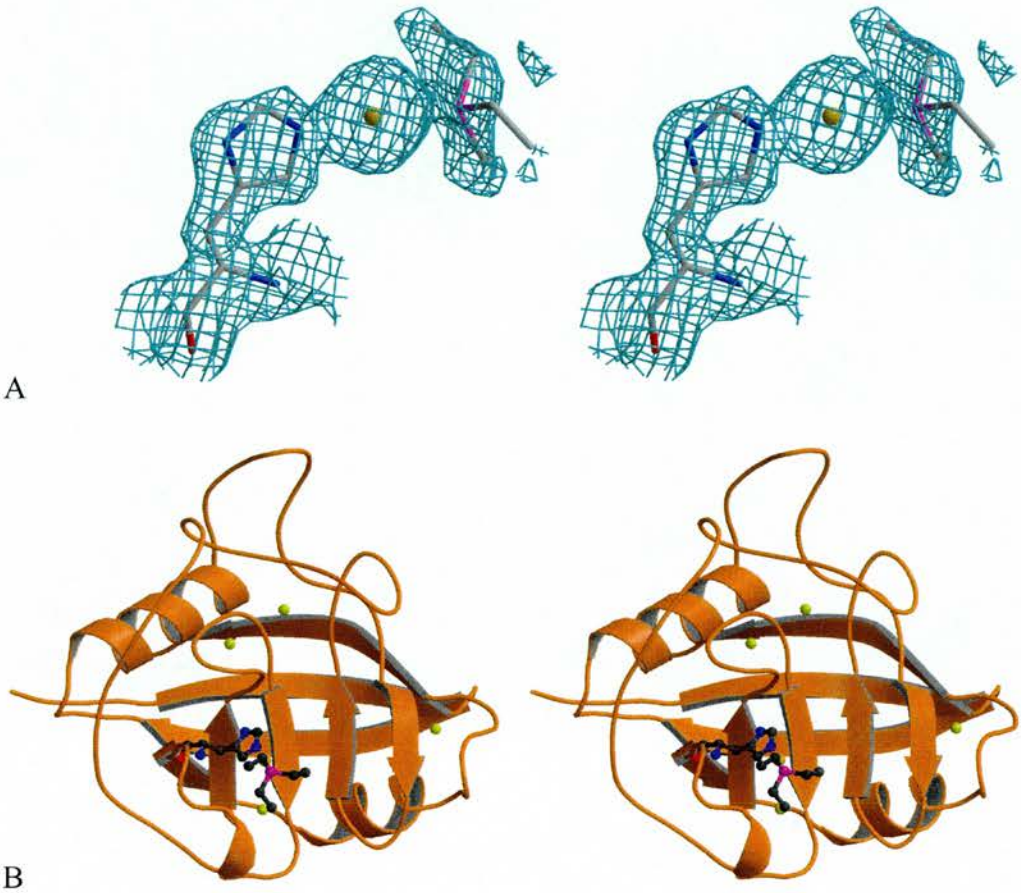


Figure 20 A and B

(A) Refined structure of the gold binding site in CYP-3, showing His 133 bound to $AuPEt_3$. Lack of definition of the PEt_3 ligand is probably due to rotational disorder.
(B) Stereoview of CYP-3 after reaction with $AuPEt_3Cl$ showing the S atoms of the 4 Cys residues (light spheres) and $[AuPEt_3]^+$ bound to the Nε2 nitrogen centre of His133. The figures were prepared using Bobscript (Esnouf, 1997), Molscript (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

These results presented above are more fully described and discussed in a publication by Zou et al., 2000.

3.7.3

Peptidyl-prolyl cis-trans isomerase activity and its inhibition by Cyclosporin A

Representative spectral progress curves showing the formation of the *trans*-isomer of the model substrate tetrapeptide can be seen in Fig 21.

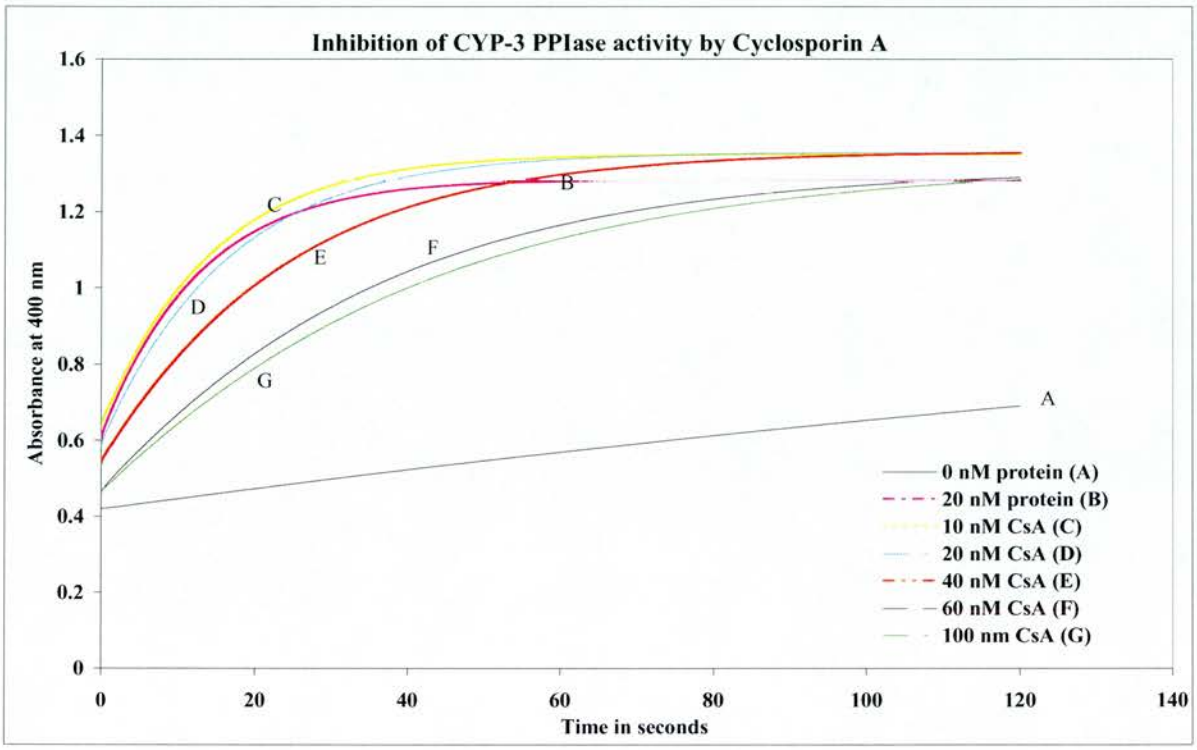


Figure 21

Representative spectral progress curves showing CYP-3 catalysed acceleration of the rate of isomerisation of the tetrapeptide substrate. The uncatalysed rate (no CYP-3) is labelled A. The curve labelled B shows the catalytic effect of 20nM CYP-3. Inhibition of the isomerisation reaction by increasing concentrations of the cyclophilin binding drug cyclosporin A is illustrated by curves C through G. The spectral curves obtained following pre-incubation with 10, 20, 40, 60 and 100nM CsA are shown.

CYP-3 clearly accelerates the rate of isomerisation of the tetrapeptide substrate relative to the uncatalysed thermal isomerisation rate and the catalysis is inhibited by addition of the cyclophilin binding drug cyclosporin A. The enzyme reaction was found to follow

Michaelis Menten kinetics with the velocity of the reaction (v) increasing with substrate concentration [S] as shown in Figure 22.

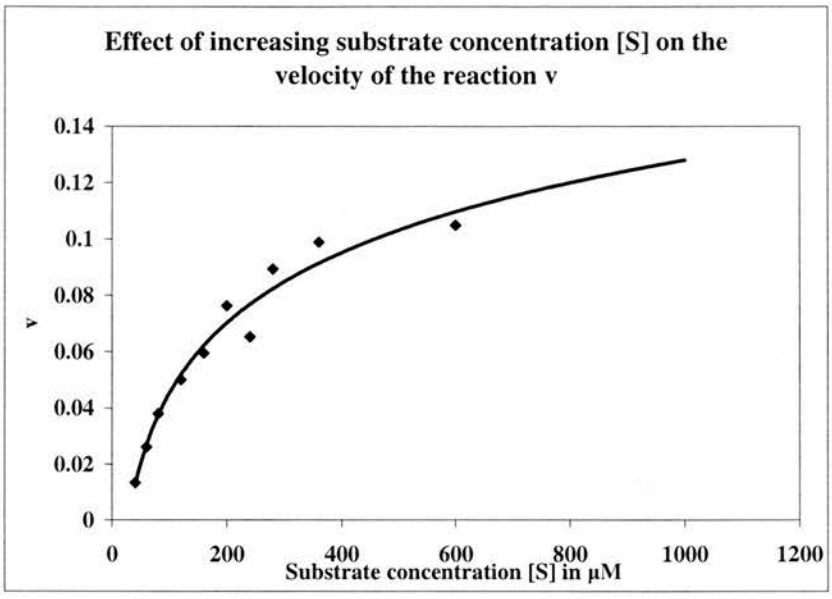


Figure 22

First order rate kinetics were observed and a double reciprocal Lineweaver-Burke plot of 1/v against 1/[S] gave values of kcat= 860 s^{-1}, Km = 350 µM which corresponds to a value of kcat/Km of 2.4 x 10^6 M^{-1} s^{-1}. These values are similar to the published values for human CypA using the same substrate which has kcat= 12700s^{-1}, Km 870 µM and kcat/Km = 14.6 x 10^6 M^{-1} s^{-1} (Kofron et al., 1991).

A concentration of 16nM CsA was required to inhibit 50% of the CYP-3 PPIase activity (IC_{50}). This value is comparable to the IC_{50} value of 19nM, obtained for the inhibition of human CypA by CsA (Liu and Walsh, 1990).

3.8

EXPERIMENTS FOCUSSING ON PEPTIDE LIGANDS BINDING TO CYP-3

3.8.1

Soaking and/or co-crystallisation experiments with CYP-3

A number of studies have shown cyclophilins to be capable of binding to a variety of peptide and non-peptidic ligands (Taylor et al., 1997) for review. CYP-3 proved to be a robust model system amenable to both co-crystallisation with and soaking in of, of a number of ligands. This was a particularly attractive model system for a number of additional reasons outlined below:

- Availability of a significant number of good quality crystals of dimensions suitable for X-ray structure analysis.
- The ability of these crystals to withstand repeated cycles of soaking followed by freezing and exposure to X-rays, allowing data for a number of points to be obtained from the same crystal.
- Short data collection times as a consequence of both the space group and point group of these protein crystals. Sufficient data can be collected overnight on the home source and within an hour at a synchrotron radiation source equipped with CCD detectors (approximately 30° of data was collected for the native structure, this data was 97% complete).
- Availability of a high resolution native structure, allowing relatively straightforward analysis of any changes induced by ligand binding.

3.8.2

The First Direct Determination of a Ligand Binding Constant in Protein Crystals

Collaborative experiments with Dr. Su-ying Wu investigating ligand binding to CYP-3 were carried out.

3.8.2.1

Ala-Pro binding to CYP-3

Soaking experiments were carried out with a number of small molecules and peptides, in an attempt to better understand the specificity of binding in this system. The di-peptide Ala-Pro provided a good model system with which to investigate the dynamics of soaking small peptide ligands into CYP-3 protein crystals.

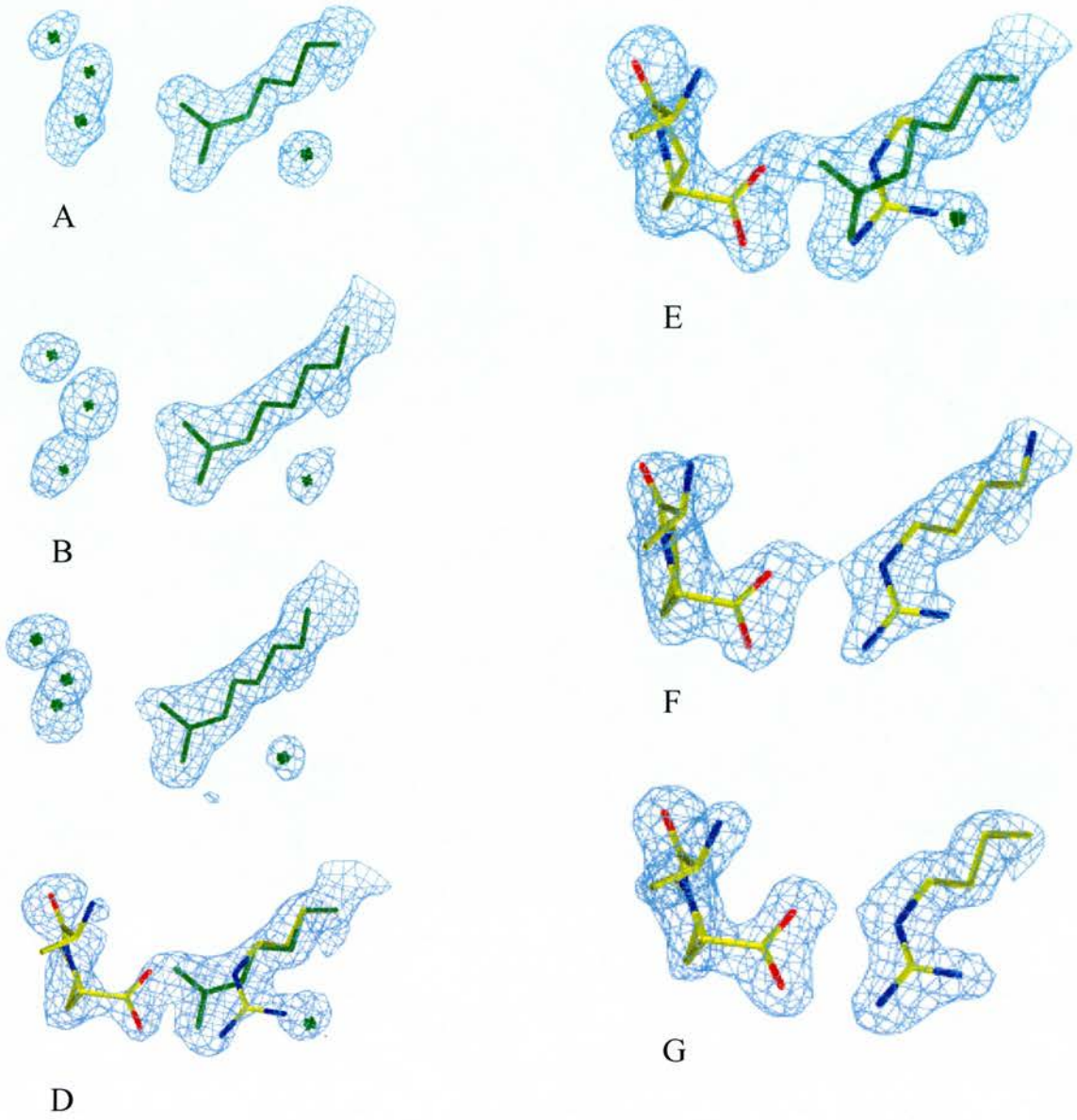
The structures of Ala-Pro bound to CYP-3 were solved and refined and the occupancy of Ala-Pro in the crystal calculated by Dr. Su-ying Wu as part of her PhD studies. This data was correlated with that obtained from studies of the protein in solution, using the PPIase activity assay. The validity of using data from protein crystal structures with ligands bound as the basis of evaluating the system in solution has long been questioned. Results presented in this study suggest that, certainly for this model system, it is both a valid and useful approach to take.

The starting point for this series of experiments were CYP-3 crystals grown as described previously, soaked overnight (between 17 and 27 hours) with Ala-Pro at a final concentration of 120mM in well solution. The crystals were harvested into freezing solution, followed by flash freezing in liquid nitrogen and exposure to X-rays. Due to their robustness, the same crystal could be used more than once, allowing collection of data for two ligand concentrations from the same crystal. Intermediate concentrations of ligand were soaked into crystals in a similar manner, downstream processing was as for the 120mM soaked crystal. All diffraction data were collected on the home source.

A series of CYP-3 crystals were soaked against different concentrations of Ala-Pro (1.2 – 120mM), and their individual X-ray structures refined. The high resolution structure of native CYP-3 was used as a template. The concentrations of ligand were as follows, 1.2mM, 6mM, 18mM, 36mM, 60mM and 120mM. This series of structures with different concentrations of ligand bound were used to determine the degree of ligand occupancy in the crystal, at the same time the binding constant was investigated in solution using the PPIase activity assay.

The assay was carried out as described in Chapter 2, Section 2.6. Protein was pre-incubated with ligand at appropriate concentrations, substrate added and the absorbance change at 400nm due to the release of p-nitroaniline was recorded, with data collected every 0.1 second over a two minute time period. Analysis of the data produced allowed calculation of a concentration of Ala-Pro (23.3mM), producing 50% inhibition of PPIase activity. This IC_{50} value was taken as an estimate of K_d in solution.

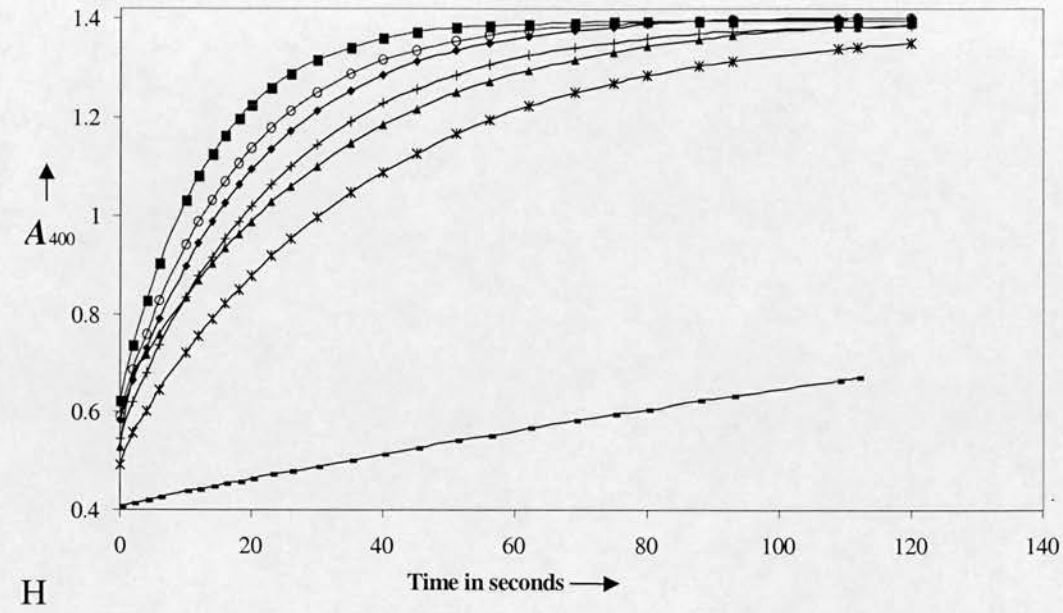
Figure 23 illustrates data obtained from both crystallographic and solution studies carried out using different concentrations of Ala-Pro. Figure 23 A-G illustrates difference electron density maps of crystals soaked with increasing ligand concentration. Figure 23 H and I illustrate representative spectral progress curves showing inhibition of CYP-3 PPIase activity with increasing concentrations of Ala-Pro and the plot of K_{obs} against inhibitor concentration respectively.



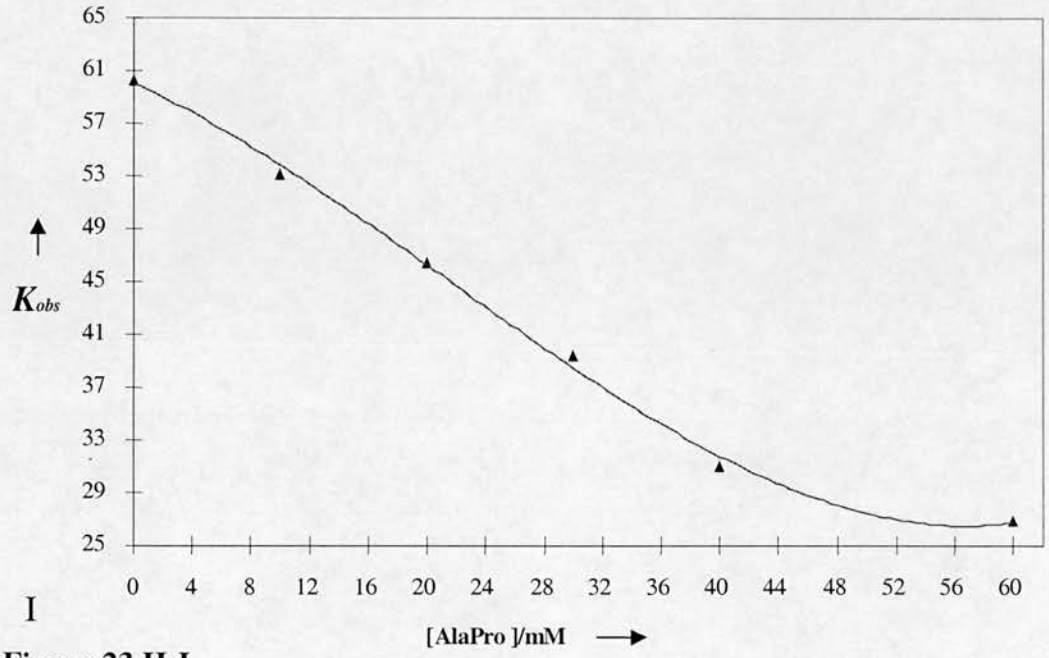
Figures 23 A - G

Difference Electron Density Maps of Crystals Soaked with Increasing Ligand Concentration

(2Fo-Fc) maps highlighting the ligand, Arg62 and active site water molecules were contoured at 1.5 sigma using the program O (Jones et al., 1991). (A) The native Cyp 3 structure shows density for the water molecules. (B and C) The crystals were soaked in 1.2 mM and 6mM Ala-Pro respectively. The electron density contoured at 1.5 sigma looks almost identical to the native structure. (D and E) The crystals were soaked in 18 mM and 36 mM Ala-Pro respectively and can be clearly interpreted as having features of both the native (green) and liganded (colour-type) forms. (F and G) The crystals were soaked in 60 and 120 mM Ala-Pro respectively. The electron density contoured at 1.5 sigma shows only features corresponding with the fully liganded form.



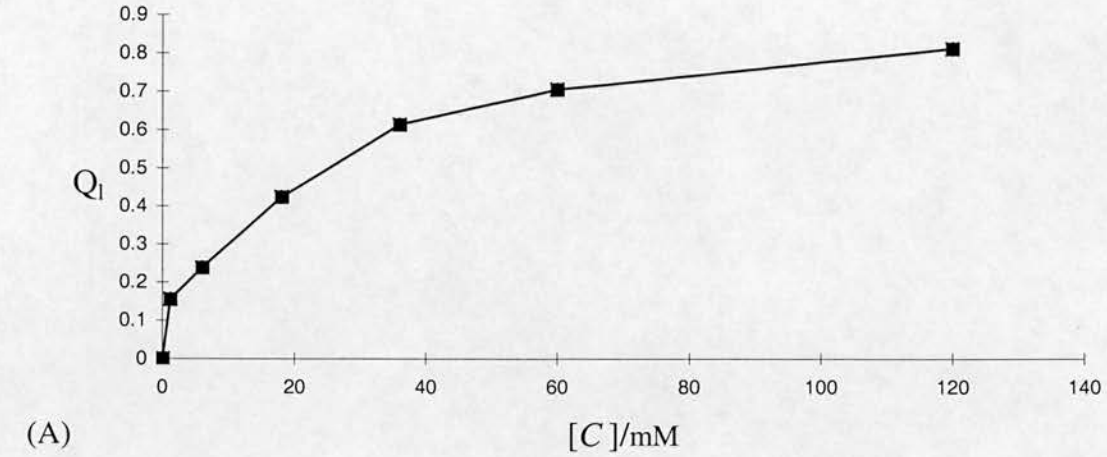
H



I

Figure 23 H-I

(H) Inhibition of PPIase activity of CYP-3 with increasing concentrations of Ala-Pro. Absorbance values were collected every 0.1s over 2 minutes. The protein concentration was 20 nM. Concentrations of Ala-Pro range between 0 to 60 mM were added. (I) Determination of the K_d for PPIase inhibition of CYP-3 by Ala-Pro. Initial velocities from Figure 23H were used to determine k_{obs} ($k_{obs}=v/[S]$) where v is the velocity and [S] is the substrate concentration (Liu and Walsh, 1990). A plot of k_{obs} against inhibitor concentration gives an Ala-Pro concentration of 23.3 mM for 50 % inhibition. This IC_{50} value can be taken as an estimate for K_d using the following assumptions: $K_d=[E][L]/[EL]$ and at 50% inhibition: $[E] = [EL] = 1/2\ [E]_T$(where $[E_T]$ = total protein concentration, [E]= free protein concentration, [L]= free ligand concentration and [EL]=protein concentration bound to ligand). With a large excess of ligand, the assumption $[L]_{50\%}=[L] \approx [L]_T$ can be made. At the point at which half maximal inhibition occurs [EL]= [E] and $K_d=[L]_T$.



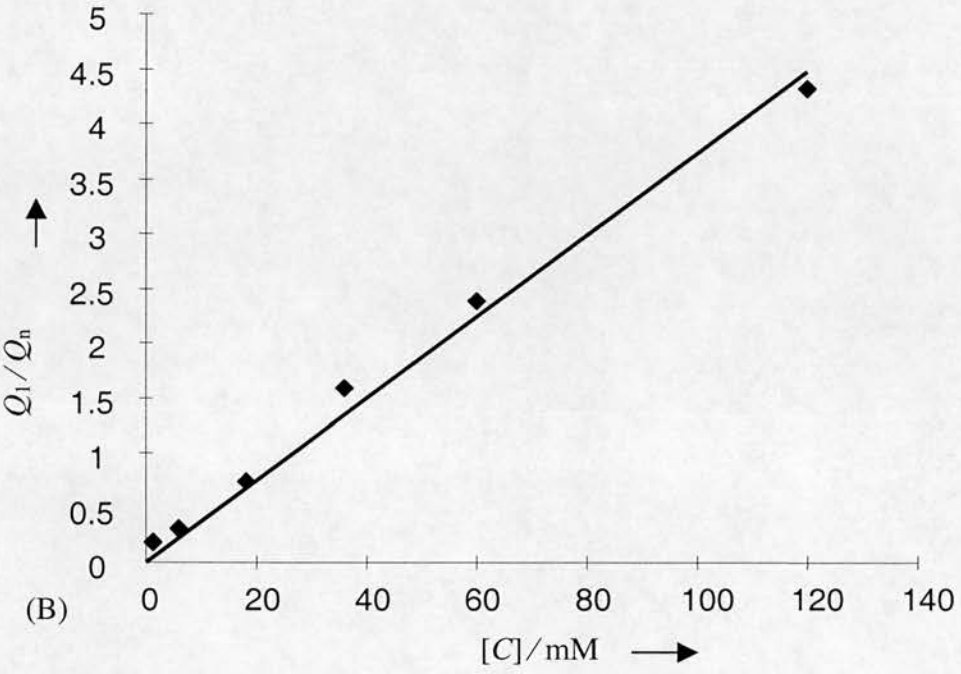


Figure 24A
Refined occupancy of the seven different crystal structures
Refined occupancy of the seven different crystal structures. Each point on the curve represents the occupancy of Ala-Pro, as determined from the refined X-ray structures of crystals soaked in a given concentration of Ala-Pro. The ligand concentrations used in the soaking experiments were 0, 1.2, 6, 18, 36, 60 and 120 mM.

Figure 24B
Plot of Ligand concentration against Q_l/Q_n
Q_l: occupancy of ligand (fraction of the protein sites in the crystal occupied by ligand)
Q_n: fraction of the protein sites in the unliganded state (= $(1-Q_l)$)
The ligand occupancies of the different crystals were measured by crystallographic refinement described in the text. The experimental data fit the theoretical line with high R^2 of 0.990 and a slope $1/K_{cd}$ = 0.0373. This gives a value for K_{cd} (the crystal dissociation constant) of 26.8 mM.

CYP-3 was a very useful template molecule for this particular series of experiments due to the relatively small change seen in the conformation of the molecule upon ligand binding. The side chain of Arg62 moves slightly to accommodate the ligand, no other significant changes are seen with this small di-peptide ligand.
To correlate solution study data with that obtained from X-ray analysis of the crystals, it was necessary to determine the degree of occupancy of ligand at each concentration used in the crystal soaking experiments.

Figure 23 illustrates electron density around the ligand and Arg62. The two conformations, free and bound, can be clearly distinguished in the density. The occupancy of each state was determined by the refinement program SHELX-97 (Sheldrick, 1997)
This series of experiments revealed a very close correlation between the results obtained from solution studies and X-ray crystallographic analysis. A plot of CYP-3 PPIase activity with increasing concentrations of Ala-Pro clearly showed the peptide having an inhibitory effect. Initial velocities were calculated from this series of peptide concentrations, and were used to determine k_{obsd} ($k_{obsd} = v/[S]$), where v is the velocity and [S] is the substrate concentration (Liu and Walsh, 1990).

This series of results is more fully described in the following publication, (Wu et al., 2001)

3.9
DISCUSSION

The structure of Cyp 3 has a number of features that distinguish it from the "classical" cyclophilin, Cyp A. In the work presented in this Chapter, Cyclophilin 3 has been characterised, as a member of the "divergent loop" subfamily of cyclophilins. The determining features of the sub-family include an insert of 7 or 8 residues forming a short loop, a highly conserved glutamate pinning down this loop in a specific fashion and two highly conserved cysteines in close proximity that remain in the reduced form. These features seen in the CYP-3 structure, when considered together may provide further sub-specialisation within the "divergent loop" sub-family of cyclophilin proteins.

The enzymatic activity of the protein was also established, and was shown to be broadly in line with that of human Cyp A, as may have been expected from complete conservation between the two proteins of residues critical in binding CsA. The binding affinity of Cyp 3 to CsA was somewhat altered, perhaps suggestive of different natural substrates for the two proteins. This also might be expected, due to the presence of 3 Cyp A type cyclophilins from *C. elegans*, suggesting a number of possible sub-specialisation's within the family.

The presence of so many cyclophilin isoforms in *C. elegans* is intriguing but also problematic, due to the increased likelihood of functional redundancy between similar isoforms of the protein. Characterisation of both temporal and spatial expression patterns may help shed more light upon when these proteins are active. A role for the PPIase activity of cyclophilins in the folding of newly synthesised collagen molecules had been suggested. The expression pattern seen for Cyp 3 peaks at the second larval stage. This expression pattern may correlate with an increase in the amount of newly synthesised structural proteins produced by the developing larvae. Defining a role for Cyp 3 in the processing of these proteins was however hard, particularly due to the highly specific expression pattern seen. The protein is exclusively expressed in the anterior excretory cell. As yet there is no evidence to support a role for Cyp 3 in any of the functions associated with the excretory cell.

The role(s) played by Cyp 3 remain to be determined. A possible role in signalling a response to oxidative stress has been suggested, due to the presence of two cysteines

in close proximity, both of which are present in the reduced form. Modelling studies suggest that a disulphide bond could form between these residues, following a simple rotation around the Cα-Cβ cysteine side chain bonds. Experiments attempting to induce such disulphide bond formation by exposing native crystals to a number of metals (attempting to simulate an oxidising environment) were largely inconclusive as the soaked crystals diffracted poorly or not at all. This was however indicative of movement within the crystal lattice, suggesting the metal soaks did have some effect upon the crystals. Experiments using various concentrations of hydrogen peroxide were also carried out, on both protein in solution and native crystals (Carnegie Summer Studentship, held by Simon Wilkinson). Solution data suggested one disulphide bond was formed, however once again soaked crystals failed to diffract, suggestive of a disruption in the crystal lattice.

The activation of proteins by disulphide bond formation has been shown previously for a number of proteins and is reviewed in (Aslund and Beckwith, 1999). Two examples will be considered in more detail. The *E. coli* transcription factor OxyR is activated by reversible disulphide bond formation and is deactivated by enzymatic reduction (Zheng et al., 1998). The small molecular chaperone Hsp 33 has also been shown to be activated by this mechanism. Under reducing conditions all six cysteine residues of Hsp 33 are present as thiols, and the protein exhibits no chaperone activity. Exposure to an oxidising environment however resulted in the formation of two intra-molecular disulphide bonds followed by activation of chaperone activity. Recent publications have reported the two stage activation and crystal structure of this small chaperone protein (Graumann et al., 2001; Vijayalakshmi et al., 2001) respectively. A second structure of Hsp 33 waas published shortly after (Kim et al., 2001). In both structures, the "chaperone active" oxidised protein is present as a domain swapped dimer, the formation of the disulphide bond bring crucial in both activation and dimerisation of the protein. These results showing induction of conformational change following activation and disulphide bond formation in Hsp 33 are broadly in agreement with a possible explanation for results from soaking experiments carried out on Cyp 3. Induction of disulphide bond formation by either metal or hydrogen peroxide catalysed oxidation may have induced significant conformational change within regions of the molecule, leading to perturbation of the crystal lattice, with consequent loss of diffraction from the crystals.

Cyp 3 proved to be a very useful and robust model system to study protein-ligand interactions. The crystals were amenable to soaking, froze well and were capable of withstanding freeze thawing allowing collection of more than one data point from a concentration series from a single crystal. Studies using the dipeptide Ala-Pro both in solution and "*in crystallo*" allowed direct comparison between the two states and showed good correlation between data obtained by these two methods. This series of experiments provided good evidence that data obtained from X-ray crystallographic ligand binding studies does indeed mirror well results obtained in solution, at least in this experimental setup.

CHAPTER 4

PURIFICATION, CHARACTERISATION AND CRYSTALLISATION OF BOVINE CYCLOPHILIN 40

4.1

INTRODUCTION

This chapter will focus on bovine Cyclophilin 40 (Cyp 40), presenting results of characterisation studies and two crystal structures of the protein, derived from very different crystal forms.

The protein to be discussed in this thesis is derived from a clone supplied by the Ratajczak laboratory, Nedlands, Western Australia.

Cyclophilin 40 was first identified by this group whilst investigating the oestrogen receptor from bovine uterus. The receptor was isolated from tissue samples using a single, highly selective affinity chromatography step. In the absence of sodium molybdate, the receptor was isolated in association with three main partner proteins; a 90kDa protein subsequently identified as Hsp 90, and two smaller components, at 40kDa and 22kDa (Ratajczak et al., 1990). The 40kDa receptor-associated protein was purified to homogeneity using preparative SDS PAGE followed by electroelution. A later paper by the same group describes the cDNA cloning and complete sequencing of the 40kDa binding protein, allowing its identification as a new member of the growing family of cyclophilins (Ratajczak et al., 1993).

4.2
MATERIALS AND METHODS

4.2.1

Coning of bovine cyclophilin 40 (Cyp 40)

All cloning of Cyp 40 was carried out by members of the Ratajczak group, Nedlands, Australia

GST-Cyp 40

Bovine cyclophilin 40 was initially cloned as a fusion protein with glutathione-S-transferase. The following information is provided for completeness. Briefly, the approach taken was to express bovine oestrogen receptor binding cyclophilin (ERBC), referred to as Cyclophilin 40 (Cyp 40) throughout this thesis, as a GST fusion protein using the prokaryotic pGEX-2T vector (Smith and Johnson, 1988). The entire open reading frame of wild type bovine Cyp 40 cDNA was amplified using specific PCR primers with built in restriction sites to facilitate cloning into pGEX-2T. A thrombin cleavage site at the GST/Cyp 40 junction allowed the isolation of recombinant Cyp 40 after thrombin digestion and subsequent GST removal, (Ratajczak et al., 1995).

Cyp 40

To produce untagged wild type bovine cyclophilin 40, a Cyp 40 cDNA template was amplified by PCR using sequence specific oligonucleotide 5′ and 3′ primers containing Ndel and BamH1 restriction enzyme sites, respectively. A TGA stop codon was placed immediately before the BamH1 site. The PCR fragment was ligated into pGEM-T (Promega) and the sequence integrity of both ends of the insert in an isolated clone was confirmed by automated sequence analysis (Applied Biosystems). An XhoI to BclI fragment excised from wild type Cyp 40 cDNA was replaced to eliminate PCR generated errors within this region. A full-length Cyp 40 fragment produced by NdeI and BamHI digestion was then cloned into the pET -11 plasmid vector. The pET-11-Cyp 40 WT expression plasmid was transformed into the *E.coli* expression host BL-21(DE3),(Carrello et al., 1999).

4.2.2

pET expression of recombinant Cyp 40 in E.coli.

The standard protocol developed throughout the course of this project, for large-scale over-expression of proteins under control of the T7 promoter was used (see Chapter 2, Section 2.3) for both GST-Cyp 40 and Cyp 40. All work was carried out either on ice using chilled solutions, in equipment pre-cooled to 4°C, or at 4°C in the cold room

4.2.2.1

GST-Cyp 40

GST-fused recombinant cyclophilin 40 (GST-CyP40 WT) (Ratajczak and Carrello, 1996) was expressed following a 4 hour induction with 0.4 mM IPTG (ICN/Flow). Cell pellets were frozen at -20°C, thawed slowly on ice and solubilised in lysis buffer (20mM Tris pH 7.8 buffer containing 150 mM NaCl, 2 mM EDTA, 5 mM DTT, 1% v/v Triton X-100 and 5 mM benzamidine). Lysozyme was added to 0.1% w/v and the cell suspension incubated on ice for 1 hour. An equal volume of distilled water at 4°C was added, the cell suspension mixed and centrifuged at 48400g for 60 minutes. The resulting supernatant was frozen at -20°C. To begin purification the supernatant was thawed slowly to keep the protein cold, followed by filtration through a series of membranes, the final step being 0.45 μm, prior to the purification steps to be outlined later in this chapter.

4.2.2.2

Cyp 40

Induction and harvesting of the non-fusion version of the protein was essentially as described above for the GST-bCyp40 WT protein. Clarified lysate from a 6 litre culture of *E.coli* expressing Cyp 40, grown in LB supplemented with ampicillin at 100μg/ml, after cell lysis and centrifugation was bound in batch mode to Q Sepharose Fast Flow. The protein and resin were incubated for 1 hour on ice with

frequent mixing, packed into a column and attached to the Gradifrac system as in Chapter 2 Section 2.3.

4.2.3

Purification of recombinant Cyp 40

4.2.3.1

GST-Cyp 40

The protein extract was purified to near homogeneity using 4 purification steps. Unless otherwise stated all chromatographic resins were obtained from Amersham Pharmacia Biotech (APB).

1. Glutathione-sepharose affinity chromatography. Protein was applied in Buffer A (50 mM Tris pH 7.5 buffer), GST-tagged protein was eluted with Buffer A containing 25mM reduced glutathione.
2. Anion exchange chromatography. Protein from glutathione-sepharose chromatography was applied to Resource Q resin in Buffer A. Elution with a linear gradient from Buffer A to Buffer A plus 1 M NaCl gave peak fractions which were pooled, concentrated, quantitated for protein concentration by the Bradford method and dialysed overnight against Buffer A.
3. Thrombin cleavage, GST fusion protein was cleaved with purified thrombin at a ratio of 1:100 for 90 minutes at room temperature. Cleaved protein was subjected to a second round of chromatography on glutathione sepharose as already described.
4. Final ion exchange chromatography step removing thrombin and trace contaminants.

Clarified supernatant obtained from a 4 litre culture of *E. coli* expressing GST-Cyp 40 grown in LB supplemented with ampicillin at 100μg/ml after cell lysis and centrifugation was bound in batch mode to 9ml glutathione sepharose 4B. This mixture was incubated on ice for 1 hour and poured into a disposable column fashioned from a 10ml syringe. The fusion protein was eluted competitively with reduced glutathione. Fractions were tested on SDS PAGE gels, those enriched for the fusion protein (66kDa) were pooled and dialysed against Buffer A to remove glutathione, the partially purified protein was then applied to anion exchange resin,

followed by elution of bound proteins by a gradient of NaCl. At this stage, the fusion protein was sufficiently pure to attempt thrombin cleavage, followed by another round of ion-exchange chromatography.

4.2.4
Removal of the GST affinity tag

GST is a large molecule of approximately 26kDa, therefore prior to crystallisation trials its removal was essential. The expression plasmid contained a thrombin cleavage site at the GST/Cyp40 junction. Thrombin is a serine protease that cleaves only certain Arg - X and less frequently Lys –X bonds with a clear preference for Pro proceeding the Arg or Lys. Two methods were employed to remove the GST tag and are outlined below.

4.2.4.1
Thrombin immobilised on Affigel 10

Human thrombin purified in house and supplied by Dr Holger Husi was coupled to Affigel 10 (Biorad) following manufacturers instructions, and packed into a disposable column, fashioned from a syringe. Clarified lysate enriched for the fusion protein following affinity chromatography was cycled repeatedly over the Affigel resin, with samples for SDS PAGE analysis taken at regular intervals. The application of protein to the resin was repeated until no intact fusion protein could be visualised on gels. The resulting material was re-applied to glutathione sepharose, removing any remaining traces of un-cleaved material and also the cleaved GST tag. To remove contaminating thrombin the protein solution was applied to Resource S resin. SDS PAGE analysis showed the resulting protein after Resource S ion exchange chromatography was largely free of contaminants, therefore the protein was dialysed against 50 mM Tris pH 7.5 and concentrated for further characterisation and crystallisation trials.

4.2.4.2

Direct addition of thrombin to the protein solution

Prior to attempting a large-scale thrombin digestion of fusion protein, it was necessary to determine the optimum enzyme to substrate ratio required for the reaction. This was achieved by incubating 4-5μg aliquots of fusion protein with increasing amounts of thrombin (stock concentration 2.3mg/ml), followed by analysis on SDS PAGE gels. The enzyme to substrate weight ratio required for efficient cleavage of pGEX-2T recombinant proteins generally ranged between 1:500 and 1:100.

Conditions for optimal cleavage of the fusion protein were determined following a number of small scale trials varying the following factors (1) ratio of purified thrombin to fusion protein, (2) time and (3) temperature. A number of other parameters had also to be considered, including minimising unwanted proteolytic activity, denaturation of the purified protein and also efficient usage of the limited amount of thrombin available.

4.2.5
Untagged Cyp 40

Purification to near homogeneity was achieved in two ion exchange steps.

1. Anion exchange chromatography. Protein was applied to a column of Q Sepharose. In Buffer A, 50mM Tris pH 7.5, the column was washed to baseline and then developed with a gradient of Buffer A plus 1 M NaCl. Protein was eluted by 150 mM NaCl and fractions were analysed on 12% SDS-PAGE gels followed by staining with Coomassie Blue. Fractions containing Cyp 40 were pooled and dialysed overnight at 4°C against 20 mM Tris pH 8.0 buffer (Buffer C).

2. The protein sample was applied to Mono Q resin in Buffer C, retained protein was eluted with a gradient of Buffer C plus 0.5 M NaCl. Positive fractions were pooled and concentrated, protein concentration was assayed by a modified version of the Bradford assay.

4.2.6

Biochemical pull-down assay to isolate proteins specifically retained by Cyp 40

In an attempt to isolate partner proteins for Cyp 40, GST-Cyp 40 was used in a preliminary "fishing" experiment. GST-Cyp 40 was incubated with *S. pombe* whole cell extract, followed by incubation with glutathione sepharose resin. Unbound proteins were removed by repeat washes with buffer. Proteins retained on the resin were subjected to SDS PAGE analysis.

4.3

CHARACTERISATION OF CYP 40

4.3.1

Dynamic light scattering analysis of the purified protein

Dynamic light scattering studies of the purified protein were carried out as described in Chapter 2, Section 2.4.1, a number of factors were varied, including protein concentration and temperature.

4.3.2

Electrospray ionisation mass spectrometry

The protein was further characterised by electrospray ionisation mass spectrometry on a Micromass Platform II spectrometer. The sample was introduced into the instrument stream via an infusion pump, the m/z spectrum deconvoluted using MaxEnt software. Thanks to Miss Violet Anderson for the mass spectrometry results.

4.3.3

Crystallisation of Cyclophilin 40

The hanging drop vapour diffusion method as described in Chapter 2, Section 2.5.1.3 was used throughout this series of experiments.

4.3.4

Data collection

Crystals were transferred from the hanging drop using an appropriately sized cryo-loop into a cryo-protectant solution containing mother liquor plus 20% glycerol. Crystals were removed from the cryoprotectant solution after 5 seconds and mounted in a cryo-loop and flash frozen in liquid nitrogen. X-ray data were collected at 100K using a MAR image plate and processed using DENZO (Otwinowski and Minor, 1993).

4.4
STRUCTURE DETERMINATION OF TWO CRYSTAL FORMS OF CYCLOPHILIN 40

4.4.1
Structure solution and refinement

Structures resulting from two different crystal forms, monoclinic and tetragonal Cyp 40, were solved and refined by Dr. Paul Taylor of the Structural Biochemistry Group, ICMB, University of Edinburgh.
The structures were solved by molecular replacement. A single domain cyclophilin was used as a model in both cases.

4.4.2
Monoclinic Cyp 40

The structure was refined using data to 1.8 Å resolution and gave a final R= 17.83%. All residues were well defined in the final electron density with the exception of the five C-terminal residues. The architecture of the complete molecule shows two distinct domains (Figure 32); an N-terminal cyclophilin domain from residues 1 to 183, a linker region of 30 amino acids from residues 184 to 213 and a C-terminal TPR domain of 157 amino acids from residues 214 to 370.

4.4.3
Tetragonal Cyp 40

The structure was refined using data to 2.2Å resolution and gave a final R=24.32%. Residues 2 – 298 of the structure are well defined in the final electron density, however the C-terminal 72 residues were not visible in the electron density.

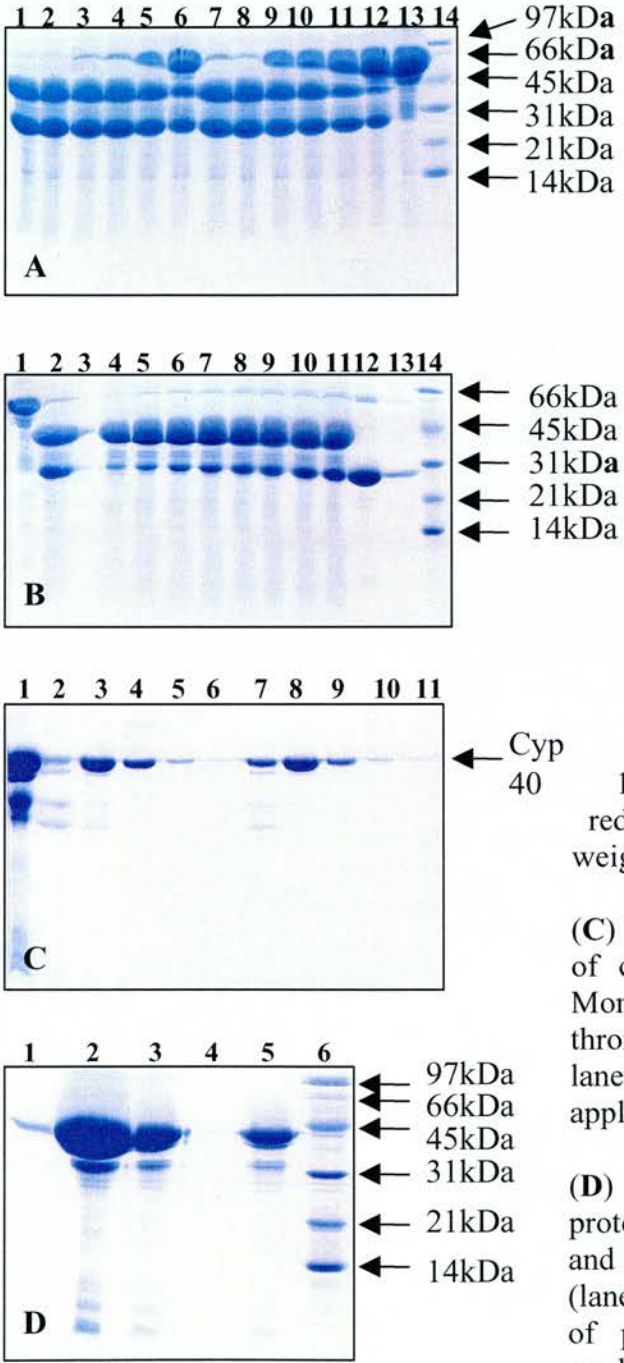
4.5

RESULTS AND DISCUSSION

4.5.1

Purification of GST-Cyp 40 and Cyp 40

4.5.1.1 GST tagged Cyp 40



(**A**) Results of thrombin cleavage trials of GST-Cyp 40. Gel illustrates experiments on the effect of temperature and time on protein cleavage. Lanes 1-6, protein incubated with thrombin at room temperature for 60, 50, 40, 30, 20, and 10 minutes respectively. Lanes 7-12, protein incubated on ice, for the same time intervals. Lane 13, GST-Cyp 40 pooled and concentrated following elution from glutathione sepharose. Lane 14, molecular weight markers.

(**B**) Gel illustrating application of cleaved GST-Cyp 40 to glutathione sepharose, to separate uncut protein and tag from cleaved material. Lane 1, uncut, lane 2, cleaved material, lanes 3-11, unbound Cyp 40, lanes 12 and 13, GST protein eluted with reduced glutathione. Lane 14, molecular weight markers.

(**C**) Illustration of "negative" purification of cleaved Cyp 40. Protein applied to Mono S resin, removing contaminants and thrombin. Lane 1, cleaved start material, lanes 2-11, wash fractions from two applications to the Mono S column.

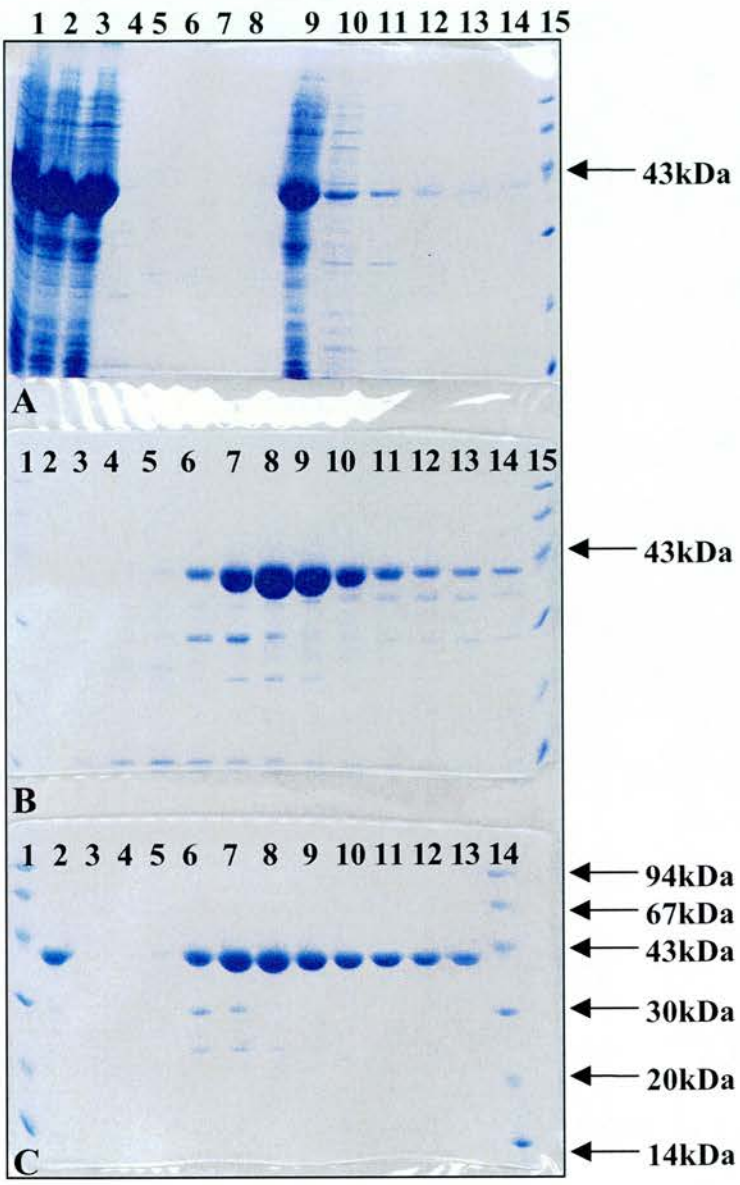
(**D**) Purified Cyp 40, Lane 1, pooled protein prior to concentration. Lanes 2, 3 and 5, 16µl, 4µl and 2µl of Cyp 40 loaded (lane 5 equivalent to approximately 16µgs of protein. Lane 6, molecular weight markers.

Figure 25

SDS page gels illustrating a representative purification run of GST-Cyp 40.

4.5.1.2

Untagged Cyp 40



(A)
A representative gel illustrating successful induction of over-expression of Cyp 40 by IPTG. Lanes 1 – 3 lysed cell pellets, lanes 4 – 6, cell supernatant. Lane 9, "start material", (clarified cell extract), lanes 10 – 14, wash fractions. Lane 15, molecular weight markers.

(B)
Representative gel illustrating a typical elution profile from Q Sepharose. Lanes 1 – 4, wash fractions. Lanes 5 – 14, elution of Cyp 40 by an NaCl gradient. The peak of protein can be clearly seen. Lane 15, molecular weight markers. Peak fractions were pooled, dialysed and applied to Mono Q resin.

(C)
Representative gel illustrating a typical elution profile from Mono Q resin. Lane 2, Cyp 40 following elution from Q Sepharose, pooled and dialysed. Lanes 3 – 13, typical elution profile from Mono Q resin. The Cyp 40 was eluted by an NaCl gradient. "pure" fractions, lanes 8 – 13 were pooled and concentrated and used in further experiments.

Figure 26

Purification of recombinant Cyp 40 was achieved in 2 anion exchange steps

4.5.2

Biochemical pull-down assay to isolate proteins specifically retained by Cyp 40

In this experiment GST-Cyp 40 was first bound to glutathione sepharose. Unbound material was eluted by washing with buffer, followed by incubation with whole cell extract of the fission yeast *Schizosaccharomyces pombe*, kindly supplied by Terri Gaskell (ICMB, Edinburgh). The resin was carefully washed a number of times to remove any unbound material. The washed resin with associated proteins was then boiled in reducing sample buffer and the protein sample subjected to SDS PAGE analysis. Two bands at approximately 85 and 45kDa were enhanced in intensity, suggestive of specific retention by GST-Cyp 40. These two bands were cut out of the gel and processed to allow sequencing of the protein within the band. The proteins within the bands were identified by sequencing as heat shock protein 90 (Hsp 90) and Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Results not shown). Thanks to Dr. Andrew Cronshaw for the protein sequencing.

Tagged and "free" Cyp 40, complementary tools to characterise Cyp 40 protein

Availability of fusion and non-fusion clones of Cyp 40 provided complementary tools to characterise the protein and its interactions.

Fusion clone

The GST fusion clone was very useful when planning and carrying out the interaction studies discussed above. Binding of the fusion protein to an immobilised support allowed the exploitation of specific bioaffinity interactions between the C-terminal TPR domain of the protein, shown to be important in the majority of interactions between Cyp 40 and its partner proteins. However, for crystallisation studies, use of the non-fusion version of the protein was favoured, for a number of reasons outlined below.

1. The best starting point for crystallisation studies is the availability of milligram quantities of pure, homogenous protein. Recombinant proteins requiring proteolytic cleavage i.e. GST tagged proteins are susceptible to the generation of heterogeneity within the protein population, due to inappropriate cleavage. Such heterogeneity is often the cause of problems in crystallisation experiments.
2. The final yield of the protein of interest may also be reduced due to failure to completely cleave the fusion tag from the protein of interest, resulting in populations of both cleaved and uncleaved material. Unlike some affinity tags; GST is a large molecule and must be removed prior to crystallisation trials.
3. Thrombin cleavage of proteins, particularly when a large amount of material is required, can be very expensive, due to the high cost of commercially available thrombin.

Non-fusion clone

The untagged clone was chosen for further crystallisation trials for the following reasons.

1. This clone over-expressed very well, producing large amounts of Cyp 40.
2. The protein was readily purified in two simple ion exchange steps, allowing rapid processing from cell pellet to purified protein.
3. The purified protein was very soluble (80mg/ml in some instances), largely homogeneous and stable at 4°C.

The above factors strongly suggested the non-fusion version of the protein was an excellent candidate for crystallisation trials, therefore all structural studies focused on crystals grown using non-fusion protein.

4.5.3

Dynamic light scattering analysis

Table 7 illustrates a typical dataset obtained from a dynamic light scattering experiment carried out using purified Cyp 40. As outlined in Chapter 2, measurements were taken at room temperature using a DynaPro 801 instrument with micro-sampler attachment.

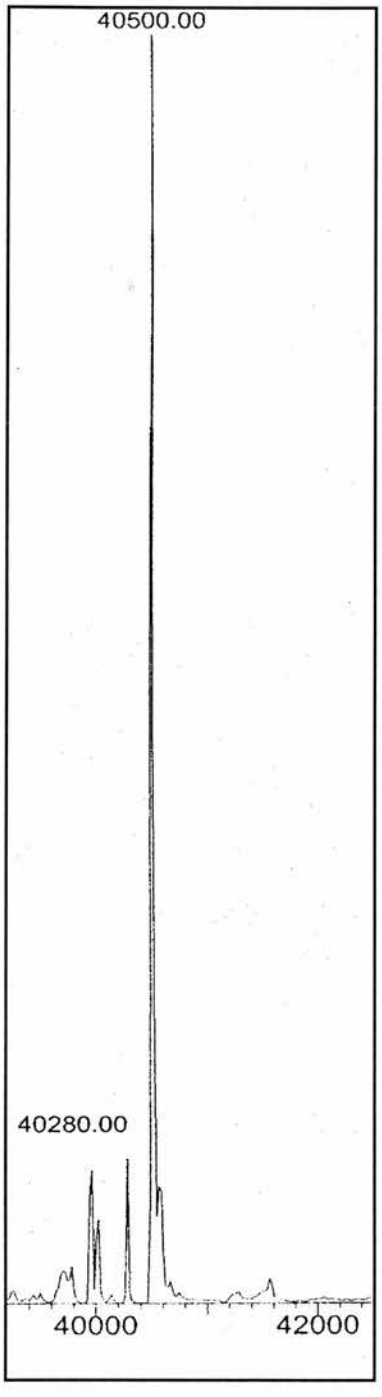
Measurement number	**Polydispersity (nm)**	**Estimated Molecular weight**	**Baseline**	**SOS Error**
1	0.9	39 kDa	1.001	2.031
2	---	41 kDa	1.000	1.284
3	---	41 kDa	0.999	1.976
4	0.9	42 kDa	1.000	2.303
5	0.8	40 kDa	1.001	1.537
6	0.7	42 kDa	0.999	1.566
7	0.9	45 kDa	0.998	2.397
8	0.7	40 kDa	1.000	2.016
9	0.9	41kDa	1.000	1.752
10	0.9	43 kDa	1.001	1.770
11	0.8	43 kDa	1.003	1.198
12	---	41 kDa	1.000	0.867
Mean values	**0.6**	**42 kDa**	**1.000**	**1.725**

Table 7

These results suggested the protein was essentially of a uniform composition (monodisperse) and present in solution as a monomer. Light scattering experiments therefore suggested the protein was an excellent candidate for crystallisation trials.

4.5.4

Mass spectrometric analysis of Cyp 40



The protein was further characterised by electrospray ionisation mass spectrometry on a Micromass Platform II spectrometer by Miss Violet Anderson. 10µl of a solution of native Cyp 40 at 80mg/ml was diluted with water to give a stock solution of 0.8mg/ml. 230µl of this stock solution was mixed with 250µl acetonitrile and 20µl of a solution of 5% formic acid. The sample was introduced into the instrument stream *via* an infusion pump at a rate of 0.48 ml/min, the m/z spectrum deconvoluted using MaxEnt software.

Figure 27

Mass spectrometry gives a molecular mass of 40,500 Da. The calculated molecular mass of the sequence is 40,620 Da and suggests that the N-terminal methionine residue is removed post-translationally. The calculated molecular mass of the sequence omitting the N-terminal methionine is 40,489 Da.

4.5.5

Crystallisation of Cyp 40

Two crystal forms were grown from the proteins isolated using the purification protocol outlined in section 4.2.5. The hanging drop vapour diffusion was used to grow crystals in 24-well Linbro plates. In all trials, crystals could only be obtained at 4°C.

The first crystals obtained were grown following refinement of conditions that produced "shaped precipitate". These preliminary crystals were typically too small to be of any use for X-ray diffraction studies, they did however provide a point from which to start. A number of rounds of refinement lead to the production of larger tetragonal crystals. The crystal morphology was generally poor, producing "teardrop" shaped crystals which typically diffracted to 3.5Å on the home source.

Alternative crystallisation conditions

Further attempts were made to produce crystals capable of diffracting to high resolution. Crystallisation trials were set up using non-fusion protein. These trials took as a starting point; conditions which had previously given small tetragonal crystals. These crystals were reproduced in the first plate set up with the non-fusion protein. Tables 8 and 9 illustrate the stages involved in obtaining useful crystals of both the monoclinic and tetragonal crystal forms.

Protein (concentration)	Buffer pH	Precipitant	Additives	Crystal morphology	Data collected	Space group	Structure solved
Cyp 40, untagged, 40 mg/ml	Cacodylate pH 6.6	560µl PEG 4000	N/A	Tiny tetragonal "teardrop" crystals	N/A	N/A	N/A
Cyp 40, untagged, 80 mg/ml	Cacodylate pH 6.2	160µl PEG 4000	N/A	Tiny needle crystals	Crystal shot on home source, no diffraction	N/A	N/A
Cyp 40, untagged, 40 mg/ml	Cacodylate pH 6.2	200µl PEG 2000	N/A	Tiny needle crystals	Crystal shot on Home source, diffracted to 4Å in one dimension only	N/A	N/A
Cyp 40, untagged, 80 mg/ml	Cacodylate pH 6.2	240µl PEG 2000	200µl Glycerol	Thin monoclinic needles	Diffracted to 3Å at SRS Daresbury	C2	Yes
Cyp 40, untagged, 80 mg/ml	MES pH 6.1	360µl PEG 4000	300µl Glycerol, 0.4M NaCl	Fatter monoclinic needles	Diffracted to 2.1Å on home source. 1.8Å at SRS Daresbury	C2	Yes
Cyp 40, untagged, 80 mg/ml	Imidazole pH 6.5	520µl MPD	20µl Glycerol, 5µl DTT 0.5mM	Large tetragonal "teardrop" crystals	Diffracted to 2.6Å on home source. 2.0Å at SRS Daresbury	$P4_222$	Yes

Table 8
Optimisation of crystallisation conditions for bovine cyclophilin 40, non-fusion clone

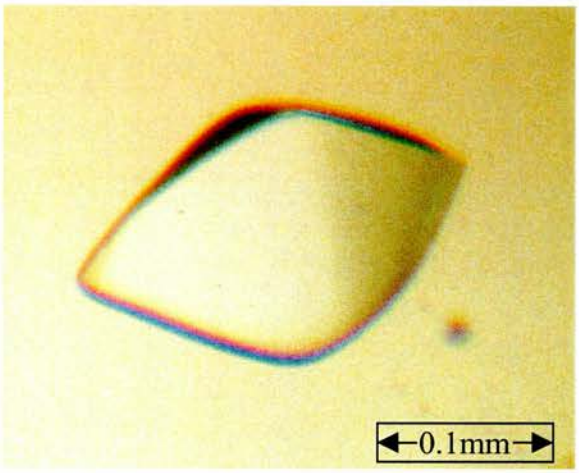
Protein (concentration)	Buffer pH	Precipitant	Additives	Crystal morphology	Data collected	Space group	Structure solved
Cyp 40, cleaved, 15mg/ml	HEPES pH 7.5	590µl MPD	N/A	"Shaped" precipitate	N/A	N/A	N/A
Cyp 40, cleaved, 15mg/ml	100mM Imidazole pH 6.5	570µl MPD	N/A	Tiny tetragonal "teardrop" crystals	Diffracted to 5Å at SRS Daresbury	N/A	N/A
Cyp 40, cleaved, 10mg/ml	100mM Imidazole pH 6.5	570µl MPD	1% Glycerol 10mM DTT	Tetragonal "teardrop" crystals	Diffracted to 3.5Å at SRS Daresbury	$P4_22$	N/A

Table 9

Optimisation of crystallisation conditions for cleaved GST-Cyp 40

4.5.6

Preliminary characterisation and data collection



A

B

Figure 28 A and B

(A) Teardrop shaped crystals with dimensions up to 0.2 X 0.15 X 0.15mm grew within a few days. These crystals typically diffracted to a resolution of 2.6Å on the home source.

(B) Monoclinic needle crystals with dimensions up to 2 X 0.01 X 0.005mm grew within hours. Occasionally, much larger plates formed with dimensions up to 0.6 X 0.2 X 0.15mm. These larger crystals typically diffracted to a resolution of 2.1Å on the home source.

4.6
CYCLOPHILIN 40 3D STRUCTURE

4.6.1

X-ray structure determination and refinement of the monoclinic crystal form of Cyp40

Two crystal forms grew with space group C2, they differed in unit cell dimensions. In most drops, very long thin plate-like needles grow within hours with typical dimensions 2 mm x 0.01 mm x 0.005 mm. These crystals were monoclinic with space group C2. Unit cell dimensions were a= 125.71 b= 47.3 c= 74.6 Å ,ß = 93.9^0. These thin needle crystals diffract typically to a resolution of 3.6 Å on a rotating anode X-ray source (see Table 8).

Occasionally much larger plates formed with dimensions 0.6 x 0.2 x 0.15 mm. The structure was solved using data from one of these larger plate crystals, all data described in this thesis was derived from these larger crystals. One such crystal was mounted in a cryo-loop and dipped in a cryo-protectant solution of mother liquor plus 24% glycerol for 5 seconds and flash frozen in liquid nitrogen.

A complete dataset was collected on a Mar 300 image plate mounted on a rotating anode generator. The crystal diffracted to a resolution of 2.1Å on the home source. This crystal belonged to space group C2 with a= 125.83 b= 47.29 c= 85.36 Å ,ß = 119.4^0. There is one molecule of Cyp40 (MW = 40,500 Da) in the asymmetric unit.

The structure was initially solved using the molecular replacement programme AMoRe (Navaza, 1994), with a single domain cyclophilin, (the structure of CYP-3 from *C. elegans* described in Chapter 3) as a model. This model gave a single 10 sigma peak in the rotation function, this rotated model gave a 9 sigma peak in the translation function. The initial model had an R factor of 55% and from the first electron density map it was possible to change residues in the cyclophilin domain to correspond to those of the Cyp 40 sequence.

Additionally, some residues of the TPR domain, mainly within helices, were visible within density, over many cycles of model fitting and refinement it was possible to manually trace almost the entire TPR domain.

Subsequent to the initial structure solution, a dataset to higher resolution (1.8Å) was collected at the SRS Daresbury. This dataset was used for the final refinement of the structure. The final refined structure had an R factor of 17.83% with an R_{free} of 25.63%.

The final refined structure runs from Ser2 to Tyr365, the 5 C-terminal residues were not visible in the electron density. All other residues are well defined in the density map, the exceptions being 2 residues in the loop region (Gly 87 and Glu 88) and Lys 172 in the C-terminal region. Data analysis indicated the structure to be of acceptable quality.

4.6.2

X-ray structure determination and refinement of the tetragonal crystal form of Cyp40

Tetragonal crystals grew with space group $P4_222$. The well solution initially consisted of 480µl MPD, 100 µl 1M imidazole at pH 6.5, 140 µl 50% v/v glycerol solution and 280 µl water. The hanging drop consisted of 2 µl well solution plus 2 µl of protein at 10 - 80 mg/ml in 20mM Tris pH 8 buffer containing 100mM NaCl. The morphology of the crystals in most drops was poorly defined, the crystals grew teardrop-shaped with dimensions of approx. 0.05 mm. Addition of 1% glycerol with increased MPD concentration (520µl) and 5-10 mM DTT was found to improve crystal size and morphology (Figure 28A), large teardrop-shaped crystals with dimensions up to 0.2 x 0.15 x 0.15 mm grew within a few days.

The unit cell dimensions were a= 94.5 Å, c= 118.3 Å. A complete data set out to a resolution of 2.6 Å has been collected on an Enraf Nonius rotating anode X-ray generator with a Mar300 image plate. A subsequent dataset collected to higher resolution (2.0Å) at the SRS Daresbury was used for the final refinement of the structure.

Initial stages of the structure determination were similar to the monoclinic form with clear rotation and translation solutions apparent. The final refined structure runs from Ser2 to Asp298. The C-terminal 72 residues were not visible in the electron density.

4.6.3

Refinement statistics

Summary of refinement statistics	Monoclinic Cyclophilin 40	Tetragonal Cyclophilin 40
Resolution range (Å)	1.8-	2.0-
R_{cryst}	17.83%	24.32%
R_{free}	25.63%	30.79%
Root mean square deviations from ideal geometry		
Bond lengths (Å)	0.07	0.06
Angle distance (Å)	0.23	0.22
Mean B-factor (Å^2)	14.90	37.63

Table 10

Summary of the final refinement statistics used in the structure determination of the two crystal forms of Cyp 40.

4.6.4

The overall architecture of bovine Cyclophilin 40

The overall architecture of Cyp 40 reveals the protein is composed of an N-terminal cyclophilin domain and a C-terminal domain comprising three tetratricopeptide repeats (TPRs) and an extended calmodulin-binding helix. The cyclophilin domain belongs to the 'divergent-loop' class with an inserted octapeptide loop and two conserved cysteine residues in close proximity to each other. A linker of 30 amino acids joins the two domains. The TPR-domain consists of seven helices of variable length incorporating three TPR-motifs. The structures of the cyclophilin domain followed by the TPR domain will be described in more detail.

4.6.4.1

The overall architecture of the cyclophilin domain of monoclinic and tetragonal Cyp 40 is conserved between the two structures

The overall architecture of the cyclophilin domain of Cyp 40 is similar to that of known X-ray and NMR cyclophilin structures (Braun et al., 1995; Taylor et al., 1997) with an anti-parallel 8-stranded β-barrel capped at each end by two α-helices. There is 61% identity with the sequence of human Cyp A (Figure 30B). The closest cyclophilin relative to Cyp 40 with known 3D structure is the divergent-loop cyclophilin from *C. elegans* (CYP-3) that shares a 64 % identity in amino acids.

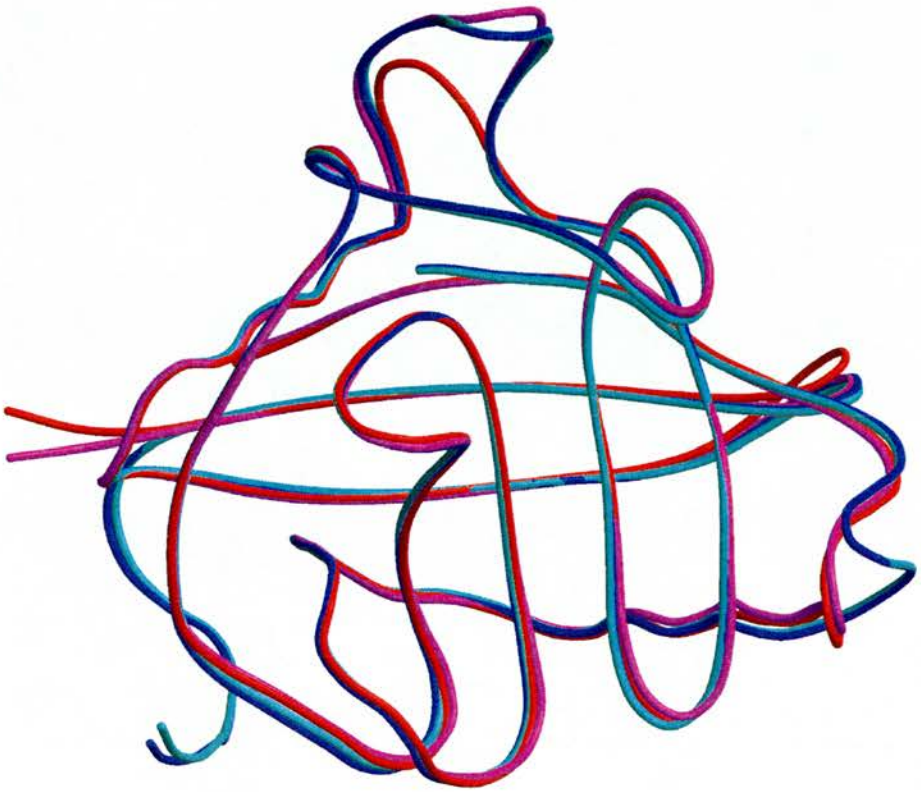


Figure 29

Overlay of the corresponding backbone atoms of the cyclophilin domains of human Cyp A (orange), nematode Cyp 3 (magenta), and bovine Cyp 40 monoclinic and tetragonal (blue/cyan). This shows that the only major difference between Cyp A and both Cyp 40 structures is caused by insertion of the 'divergent loop'. Loop regions do show some divergence as can be seen with the loop structures of Cyp 3 and Cyp 40. The figure was prepared using Molscript (Kraulis, 1991).

4.6.5

Key structural features of Cyclophilin 40

4.6.5.1

The active site of Cyp 40

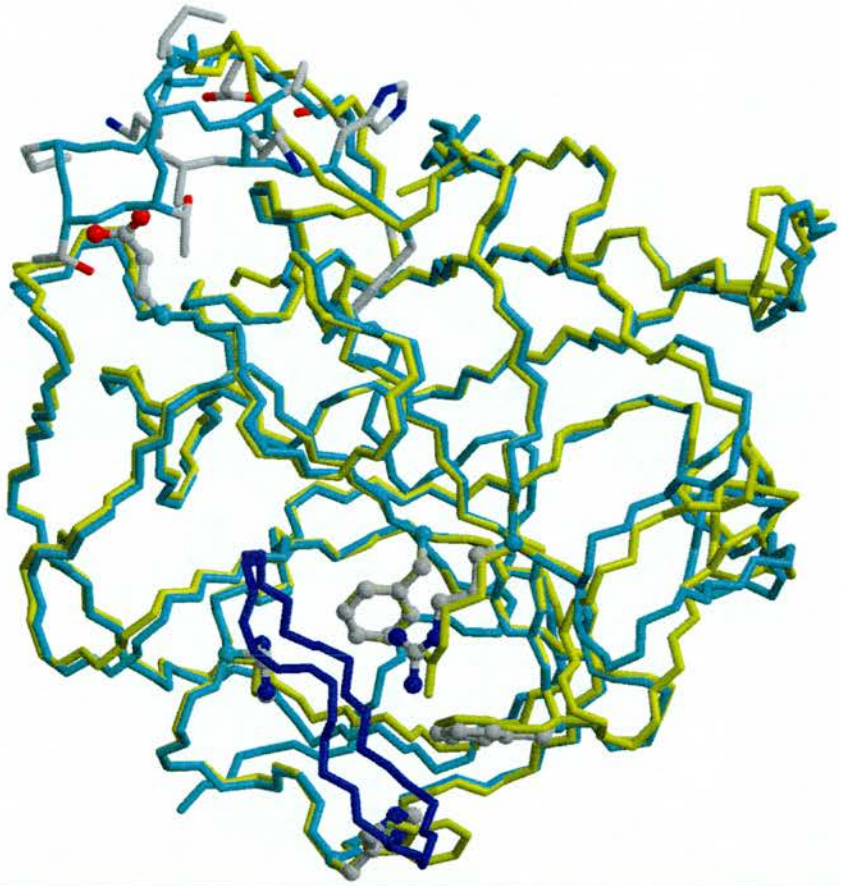
The peptidyl-prolyl cis-trans isomerase active site of Cyp40 is similar to that of human Cyp A. The active site of the cyclophilin family is defined by the 13 amino acids found to be important in the binding of CsA to human Cyp A, a number of these residues are highlighted in Figure 30A. The active site residues are identical apart from His 141 that corresponds with Trp120 in Cyp A. Site directed mutagenesis of this residue in Cyp A has shown that it plays a significant role in CsA binding, but is not so important in the PPIase enzymatic activity (Liu et al., 1991b). This observation is in line with the catalytic and CsA binding properties of bovine Cyp40.

4.6.5.2

Peptidyl prolyl cis-trans isomerase activity and its inhibition by Cyclosporin A

The PPIase activity of Cyp 40 has a kcat/Km value of 1.9×10^6 $M^{-1}s^{-1}$ (Kieffer et al., 1992) and 5.1×10^5 for GST-Cyp 40 (Ratajczak et al., 1995) compared with the value of 1.4×10^7 $M^{-1}s^{-1}$ for Cyp A (Liu et al., 1990). The inhibition by CsA is however significantly less than Cyp A with measured IC_{50} values of 300nM for free Cyp40 (Kieffer et al., 1992) and 1030nM for a GST-Cyp40 protein (Ratajczak et al., 1995). Both values suggest much weaker binding compared with the IC_{50} value for Cyp A of 19nM (Liu et al., 1990). Reduced binding affinity to CsA has been observed in a number of cyclophilins in which Trp120 of Cyp A has been changed.

A



B

```
CYP4_BOVIN MSHPSPQAKP SNPSNPRVFF DVDIGGERVG RIVLELFADI VPKTAENFRA 50
CYPH_HUMAN ---------- --MVNPTVFF DIAVDGEPLG RVSFELFADK VPKTAENFRA 38

CYP4_BOVIN LCTGEKGIGP TTGKPLHFKG CPFHRIIKKF MIQGGDFSNQ NGTGGESIYG 100
CYPH_HUMAN LSTGEKGFG- -------YKG SCFHRIIPGF MCQGGDFTRH NGTGGKSIYG 80

CYP4_BOVIN EKFEDENFHY KHDKEGLLSM ANAGSNTNGS QFFITTVPTP HLDGKHVVFG 150
CYPH_HUMAN EKFEDENFIL KHTGPGILSM ANAGPNTNGS QFFICTAKTE WLDGKHVVFG 130

CYP4_BOVIN QVIKGMGVAK ILENVEVKGE KPAKLCVIAE CGELKEGDDW GIFPKDGSGD 200
CYPH_HUMAN KVKEGMNIVE AMERFGSRNG KTSKKITIAD CGQLE----- ---------- 165
```

Figure 30 A and B

(A) Comparison of the cyclophilin domain of bovine Cyp40 (green) with human Cyp A (yellow). C, N, and Cα atoms are shown. The backbone atoms of Cyclosporin A, present in the human Cyp A crystal structure, are shown in blue. Side chains of the Cyp 40 active site residues are shown: Arg75, Phe80, Phe133, His141 and His146. The corresponding side chains for human Cyp A are shown in red. Side chains for Cyp 40 are drawn from Cys52 to His68 that incorporates "divergent loop" residues (60PTTGKPLH67). Glu96 and the two cysteines Cys52 and Cys181 are drawn as balls and sticks with atom-type colouring (oxygen coloured red, nitrogen coloured blue and sulphur coloured yellow). (B) Sequence alignment of the Cyp domain of both proteins, conserved binding site residues are highlighted in blue.

4.6.5.3

The cyclophilin domain of Cyp 40 belongs to the 'divergent-loop' family.

The 'divergent-loop cyclophilins' are determined by three features shared by the cyclophilin domain of Cyp 40. Firstly, there is an additional loop incorporating residues 60 through 67 (60PTTGKPLH67), secondly there are often two cysteine residues (Cys52 and Cys181) which are in close proximity but remain in the un-oxidised form, and thirdly there is a highly conserved glutamate residue (Glu96) that helps fix the conformation of the additional loop.

The 'divergent-loop' of bovine Cyp40 consists of an insert of eight residues 60PTTGKPLH67. The presence of inserts of between 6 and 8 residues has been described previously in a number of plant (Chou and Gasser, 1997) and nematode (Page et al., 1996) cyclophilins. A consensus sequence for this additional loop based on multiple sequence alignments is (**GK*LH) (Dornan et al., 1999a). The additional residues form a protruding loop located above the active site (Figure 29), as also observed in the other two available 'divergent-loop' cyclophilin X-ray structures of Bm-CYP-1 (Taylor et al., 1998; Mikol et al., 1998) from *Brugia malayi*, and CYP-3 (Dornan et al., 1999a) from *C. elegans*. The loop conformations adopted by all three structures are significantly different from each other. Cyp 3 forms a conventional four-residue β-turn involving 47GKSG50 while both Cyp 40 (and bmCYP-1) form unusual five-residue loops 59GPTTG63 (and 50GKISG54) which are defined by the G59O---G63N = 2.74 Å (and G50O—G54N = 2.66 Å) hydrogen bonds. Many cyclophilin homologues belonging to the 'divergent-loop' family (Dornan et al., 1999a) also have a conserved glutamate (Glu96) in a homologous position (Figure 14). The glutamate side chain ties down the loop using an interesting hydrogen bonding motif in which the two carboxyl oxygen atoms form hydrogen bonds with backbone nitrogen atoms from adjacent residues: E96OE1--T61N = 2.99Å and E96OE2--T62N = 3.10Å. It has already been suggested that the loop region may provide a trigger mechanism or recognition site in protein-protein interactions.

4.6.5.4

Conserved cysteines

Like other structurally characterised 'divergent loop' cyclophilins, Cyp 40 has two cysteines in close proximity in the reduced form. The side chains of Cys52 and Cys181 are well defined in the electron density and show a sulphur—sulphur distance of 5.4 Å. The conformation adopted by these conserved cysteine side chains is very similar to that of the other 'divergent-loop' cyclophilin structures (Taylor et al., 1998; Dornan et al., 1999a). As with these structures, molecular modelling shows that a simple rotation about the Cα-Cβ cysteine side chain bonds enables a disulphide bond to form without any significant steric hindrance. The formation of a disulphide bond in an oxidising environment may induce conformational change and provide a signalling mechanism in response to oxidative stress.

4.6.5.5

The TPR domain is attached to the cyclophilin domain by a linker of 30 amino acids.

The acidic linker region contains 11 aspartate and glutamate residues and forms a well defined hydrogen bonded structure with two β-turns and a larger loop comprising 194PKDGSGD200. The first β-turn (K185-E186-G187-D188) defines the start of the linker and a second β-turn (F205-P206-E207-D208) occurs at the centre of the deep cleft between the Cyp and TPR domains (Figure 32). The side-chain of E-207 points into the cleft and is incorporated into an intricate hydrogen-bonded network that also involves an ordered glycerol molecule. There is no direct contact between the helices of the TPR domain and the cyclophilin domain. Helices P and Q form an intimate interface to the linker. The positively charged surface of helix Q [K244, K245, K248, R251] forms salt bridges with linker residues D204 and D200 (D204/OD...K248/N =2.99Å, D200/OD2...K245/N= 2.73Å, D200/OD1...K248/N= 3.16 Å, Figure 32). The conformation of the cyclophilin domain is almost unaffected by the presence of the linker region despite the formation of numerous van der Waals contacts and hydrogen bonds to the linker and TPR domains.

4.7

THE TETRATRICOPEPTIDE REPEAT (TPR) MOTIF IN CYP 40 AND LARGER IMMUNOPHILINS

The Hsp 90 binding immunophilins Cyp 40, FKBP 52 and 51 were identified in 1993, 1994 and 1995 respectively as TPR containing proteins (Ratajczak et al., 1993; Radanyi et al., 1994; Baughman et al., 1995). The structure of Cyp 40 (whole molecule), focussing on the TPR domain of Cyp 40, will now be discussed in greater detail

4.7.1

The C-terminal domain of Cyp 40 consists of a tetratricopeptide repeat domain

The overall architecture of Cyp 40 was described in Chapter 4, section 4.6.4. The structure revealed a clear separation between the cyclophilin and TPR domains, linked by an acidic region of 30 amino acids.

4.7.1.1

The structure of the TPR domain.

The structure of the TPR domain consists largely of seven helices. The helices delineated using the program DSSP (Kabsch and Sander, 1983) form an extended helix-turn-helix pattern (Figure 32) and are labelled P,Q,R,S,T,U and V, (see Table 11). Each 34 amino acid TPR repeat consists of a helix-turn helix motif in which the helical sub-domains are referred to generically as (A_i,,B_i) and amino acids are numbered from 1 to 34 (Figure 31). The first three helices in the domain (P,Q,R) are longer than expected for TPR sequences. Helix R incorporates a proline residue (P270) that also distorts the helical conformation of the four preceding residues. (A similar distortion occurs in all HEAT repeat helices that contain a proline in almost all known B-helices (Groves et al., 1999).

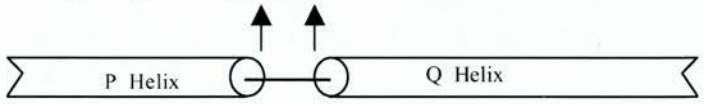
4.7.1.2

Sequence comparisons between TPR containing proteins

Figure 31 provides a comparison between the amino acid sequences of the C-terminal TPR-containing domains of these immunophilins from mammals (Cyp 40, FKBP51, FKBP52), *S. cerevisiae* (Cpr6, Cpr7) and *S. pombe*, (Wis2) with other Hsp 90 binding proteins, mammalian PP5 and CNS1 from *S. cerevisiae*. Over this entire C-terminal region Cyp40 shares a 37% identity with Cpr7, but has an increased resemblance to Wis2 (46%), FKBP52 (46%) and Cpr6 (50%). FKBP52 is most similar to FKBP51 (61%) with a lower level of homology with Cpr7 (36%), Cpr6 (40%) and Wis2 (41%).

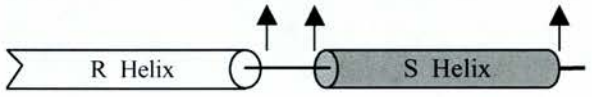
TPR 1 motifs of selected Hsp90 binding proteins

		10	20	30	34
CYP4_BOVIN	223-256	SEDLKNIGNT	FFKSQNWEMA	IKKYTKVLRY	VEGS
CYP4_HUMAN	223-256	TEDLKNIGNT	FFKSQNWEMA	IKKYAEVLRY	VDSS
FKBP5_HUMAN	268-301	AAIVKEKGTV	YFKGGKYMQA	VIQYGKIVSW	LEME
FKBP4_HUMAN	270-303	STIVKERGTV	YFKEGKYKQA	LLQYKKIVSW	LEYE
PPP5_HUMAN	28- 61	AEELKTQAND	YFKAKDYENA	IKFYSQAIEL	NPSN
CYP6_YEAST	219-252	IETVKNIGTE	QFKKQNYSVA	LEKYVKCDKF	LKEY
CYP7_YEAST	240-273	ANIIKESGTL	LFKKKDYSNA	FFKYRKSLNY	INEY
CNS1_YEAST	83-116	AENFKKQGNE	LYKAKRFKDA	RELYSKGLAV	ECED
IEFS_HUMAN	225-258	ALKEKELGND	AYKKKDFDTA	LKHYDKAKEL	DPTN



TPR 2 motifs of selected Hsp90 binding proteins.

		10	20	30	34
CYP4_BOVIN	273-306	LSCVLNIGAC	KLKMSDWQGA	VDSCLEALEI	DPSN
CYP4_HUMAN	273-306	LSCVLNIGAC	KLKMSNWQGA	IDSCLEALEL	DPSN
FKBP5_HUMAN	317-350	LAAFLNLAMC	YLKLREYTKA	VECCDKALGL	DSAN
FKBP4_HUMAN	319-352	LASHLNLAMC	HLKLQAFSAA	IESCNKALEL	DSNN
PPP5_HUMAN	62- 95	AIYYGNRSLA	YLRTECYGYA	LGDATRAIEL	DKKY
CYP6_YEAST	270-303	VSIPLNIAIC	ALKLKDYKQV	LVASSEVLYA	EAAD
CYP7_YEAST	292-325	MKIYLNLSLV	LFNLERYDDA	IMYATYLLEM	DNVP
CNS1_YEAST	121-154	ESLYANRAAC	ELELKNYRRC	IEDCSKALTI	NPKN
IEFS_HUMAN	300-333	MTYITNQAAV	YFEKGDYNKC	RELCEKAIEV	GREN



TPR 3 motifs of selected Hsp90 binding proteins.

		10	20	30	34
CYP4_BOVIN	307-340	TKALYRRAQG	WQGLKEYDQA	LADLKKAQEI	APED
CYP4_HUMAN	307-340	TKALYRRAQG	WQGLKEYDQA	LADLKKAQGI	APED
FKBP5_HUMAN	351-384	EKGLYRRGEA	QLLMNEFESA	KGDFEKVLEV	NPQN
FKBP4_HUMAN	353-386	EKGLFRRGEA	HLAVNDFELA	RADFQKVLQL	YPNN
PPP5_HUMAN	96-129	IKGYYRRAAS	NMALGKFRAA	LRDYETVVKV	KPHD
CYP6_YEAST	308-341	AKALYRRGLA	YYHVNDTDMA	LNDLEMATTF	QPND
CYP7_YEAST	330-363	AKAYYRRGNS	YLKKKRLDEA	LQDYIFCKEK	NPDD
CNS1_YEAST	155-188	VKCYYRTSKA	FFQLNKLEEA	KSAATFANQR	IDPE
IEFS_HUMAN	259-292	AKAYARIGNS	YFKEEKYKDA	IHFYNKSLAE	HRTP

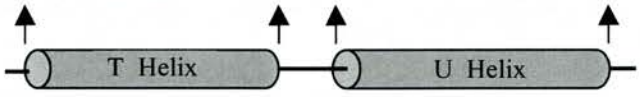


Figure 31

Alignment of TPR sequences for selected Hsp90 binding proteins.

Proteins that contain three copies only of the TPR motif and that have been shown to interact with Hsp90 were identified from the literature and are presented above manually aligned. Additionally the TPR2 domain of Hop (IEFS_HUMAN) was included. The horizontal cylinders below each alignment designate the extent of the seven helices of Cyp 40, assigned using structural data. A broken cylinder indicates the helix begins or terminates earlier or later in the protein sequence, therefore failing to conform to the classical idealised helix (13aa) turn (3aa) helix (14aa) model for TPR motifs. Arrows projecting from the cylinders point to the positions in the sequence where the helices start and finish.

Closed and shaded cylinders indicate a complete helix within the TPR motif. Only TPR 3 conforms to the idealised structure. Residues identified as being important in binding Hsp90 are red and underlined. Structurally conserved residues are italicised and blue. The three linker residues between helix A_1 and B_1 are underlined.

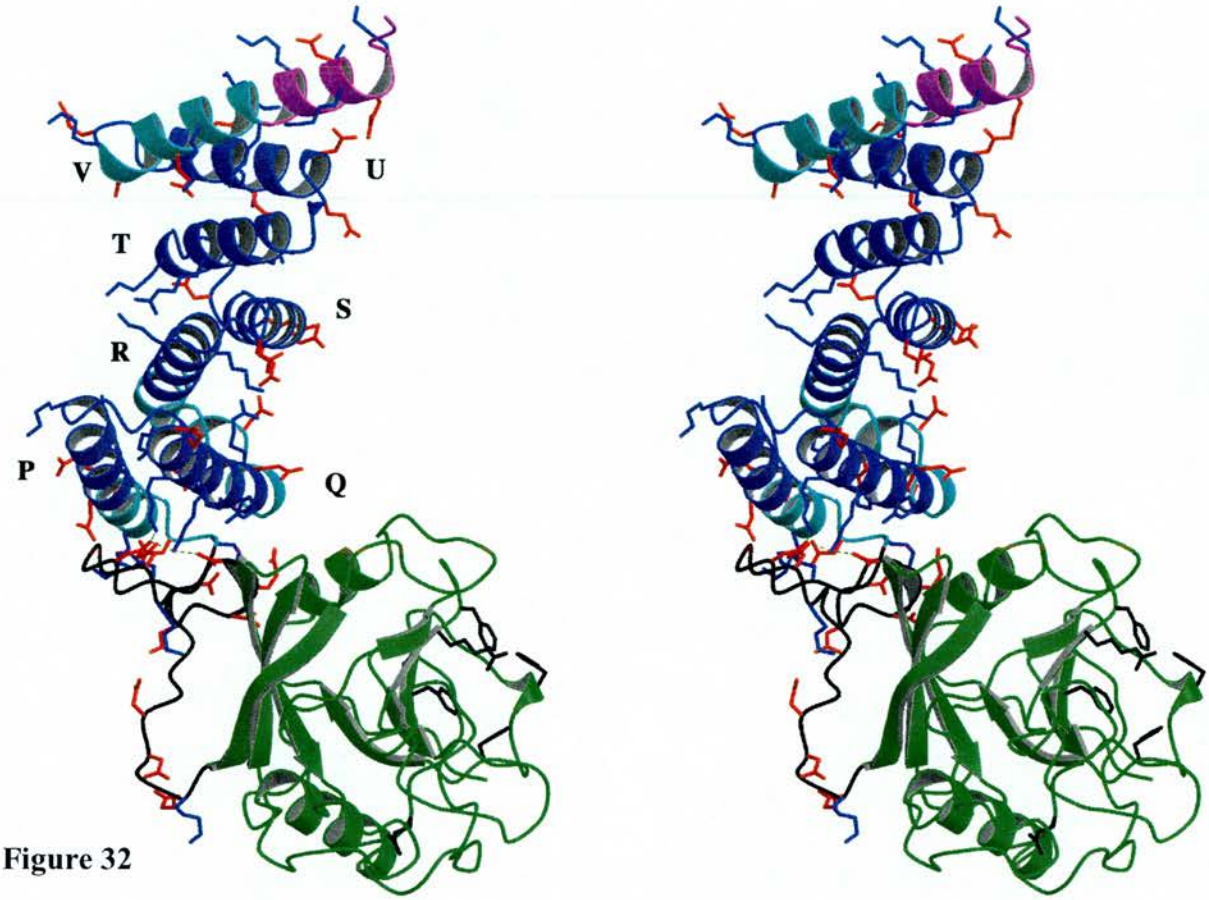


Figure 32

Stereo picture of the monoclinic form of bovine cyclophilin 40.

The cyclophilin domain (residues 1-183) is coloured green. Side chains of selected active site residues (Arg75, Phe80, Phe133, His141, His146) are coloured black. Glu96, which forms hydrogen bonds to the divergent loop, is also shown. The backbone of the linker (residues 184-213) joining the cyclophilin and TPR domains is shown black. Selected salt bridges between the positively charged residues of helix Q (K244, K245, K248, R251) with linker residues D204 and D200 are shown as dashed green lines. Side chains of all charged residues on the TPR domain are shown. Arg and Lys are blue, Glu and Asp red.

The seven helices in the TPR domain are designated P,Q,R,S,T,U,V, (see Table 11 below). The slight distortion in helix R is caused by the presence of Pro270. The portions of the helices corresponding to the repeating TPR motifs shown in Figure 31 are coloured dark blue (TPR1, helices P and Q, residues 223-256; TPR2, helices R and S, residues 273-306; TPR3, helices T and U, residues 307-340). The protruding C-terminal portion of helix V is the putative calmodulin binding domain and is coloured magenta. The figure was prepared using Molscript (Kraulis, 1991).

The seven helices of the TPR domain	Corresponding residues in protein sequence
P (20 residues)	V216 – K235
Q (21 residues)	W239 – A259
R (24 residues)	D262 – K285
S (12 residues)	W289 – L300
T (13 residues)	T307 – G319
U (14 residues)	Y323 – I336
V (22 residues)	K341 – K362

Table 11

Residues in the sequence of Cyp 40 forming the seven helices seen in the fully folded structure of the TPR domain

4.7.2

A partially folded TPR domain is trapped in the tetragonal crystal form.

The structure of the tetragonal form retains only the first TPR motif (helices P and Q); the second TPR motif (helices R and S) has sprung out to form an elongated helix of approximately 42Å, (Figure 33). Residues comprising helices T, U and V in the monoclinic form are not visible in the electron density. This structure can therefore be regarded as a trapped intermediate in the folding pathway of the helix domain. It also suggests that such helical domains are flexible and such helix extensions towards the end of a helical domain may be required in protein-protein recognition processes.

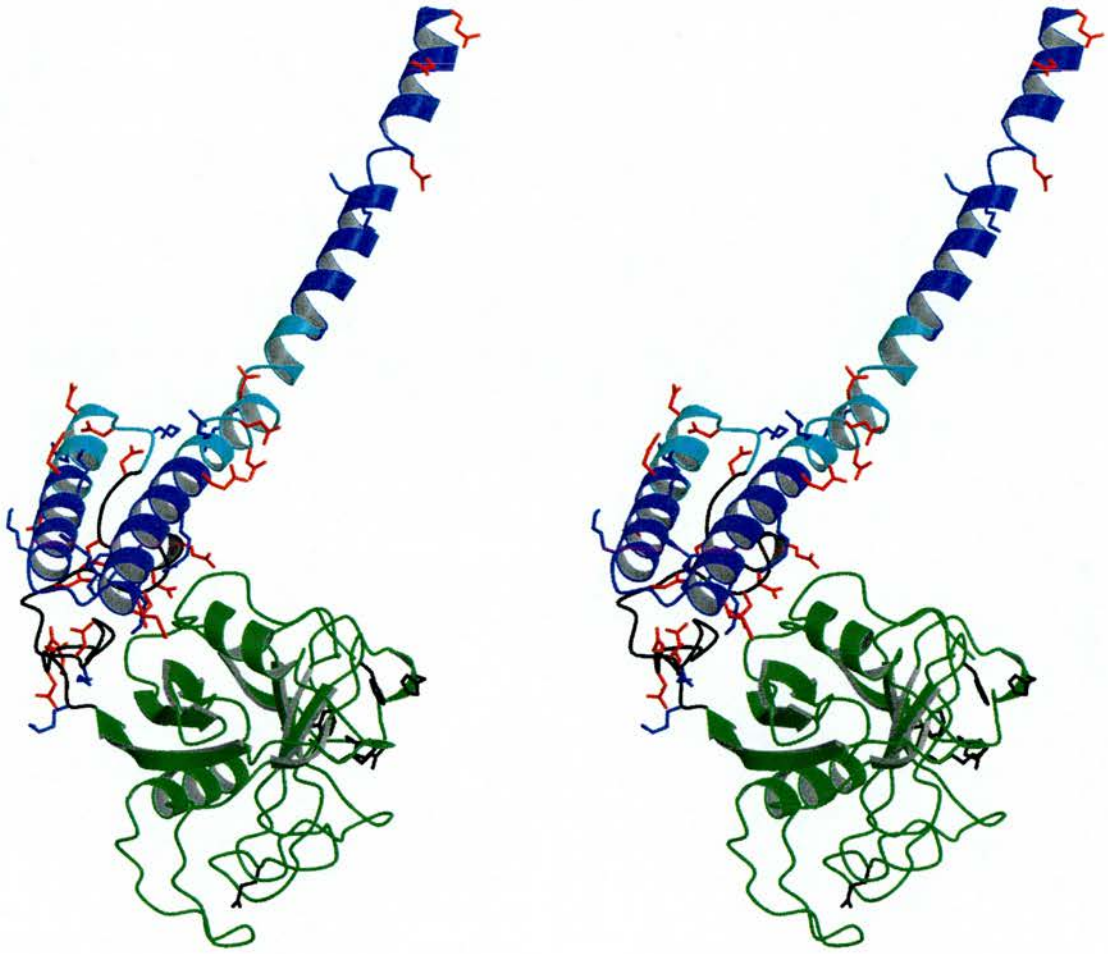


Figure 33

Stereo picture of the tetragonal folding intermediate form of bovine cyclophilin 40. Colouring is the same as in figure 32. Helices T, U and V are not visible in the electron density. Helices are coloured according to the fully folded monoclinic form in Figure 32, numbering of helices as highlighted in Table 11. The portions of the helices corresponding to the repeating TPR motifs shown in Figure 31 are coloured dark blue (TPR1, residues 223-256; TPR2, residues 273-298). Alternative conformations of residues Ala260 and Ser287 are primarily responsible for helices R and S adopting an extended rod like helical conformation. The figure was prepared using Molscript (Kraulis, 1991).

4.7.3

High fidelity recognition between two Cyp 40 structures

4.7.3.1

Tetragonal Cyp 40 - A 3D domain swapped dimer?

The conservation of interactions between the two structures raises the possibility of 3D domain swapping in the TPR domain of Cyp 40.

3D domain swapping involves one domain of a multi-domain protein breaking non-covalent bonds with the other domain(s), its place being taken by the same domain of an identical protein chain, resulting in an intertwined dimer/higher oligomer. The swapped domain can be anything from an entire domain to a particular element of secondary structure. In this example an α helix is involved in inducing dimerisation. The helices (P and Q) in the partially folded dimer, the "swapped domain", (Figure 34 A and B) make identical interactions/non-covalent bonds with the neighbouring domain (helix R) in the other half of the dimer and also with those of the monomer. This results in stabilisation of the dimer by the same interactions responsible for holding the monomeric helical domain together.

3D domain swapping has been suggested as a mechanism of protein polymerisation leading, in some cases, to the production of fibrils typically found to accumulate in amyloid plaques (Liu et al., 2001; Janowski et al., 2001).

A role for 3D domain swapping in accumulation of folding intermediates was also suggested, with folded but otherwise unassociated domains "domain swapping", trapping the folding intermediate in an oligomeric state. The tetragonal crystal form may be an example of such a phenomenon (Bennett et al., 1995).

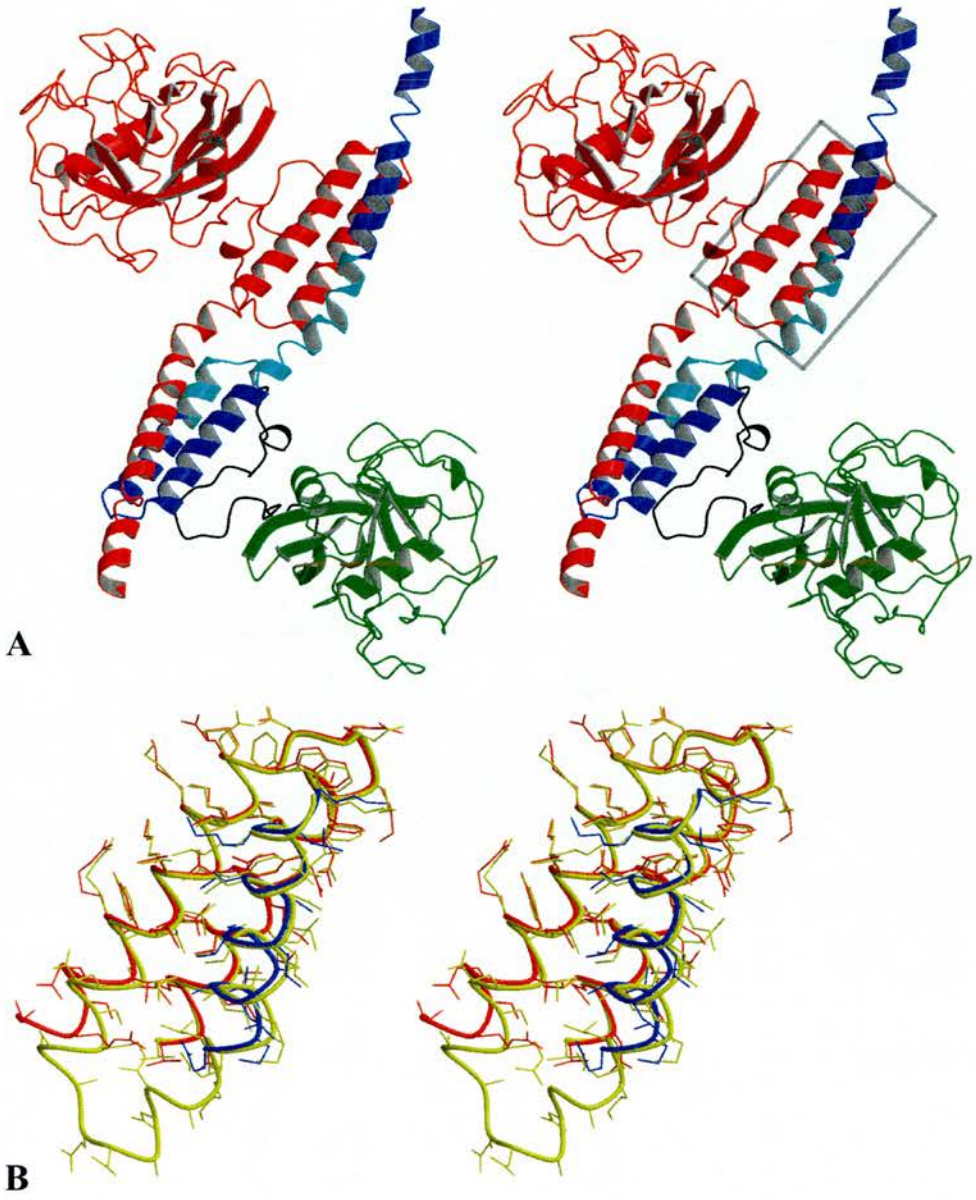


Figure 34 A and B

Stereo picture of the main intermolecular interactions in the tetragonal form of Cyp 40 and an overlay of the folding intermediate dimer and the fully folded confomer.

(A) The partially unfolded TPR domain forms a symmetrical dimer. One molecule is coloured as in Figure 32, the partner molecule is coloured red. The first TPR comprising helices P and Q (shown in the boxed region), make an intermolecular contact with helix R' of the dimer-related molecule.

(B) The boxed region of (A), with the same colour scheme, showing helices P and Q and R' for the tetragonal form, is overlaid with the helices P, Q and R (yellow) of the folded TPR domain of the monoclinic form. The intra- and intermolecular interactions of helix R with helices P and Q can be seen to be nearly identical.

4.7.4

Amino acids in the TPR motif that control helix packing

The average Ai-Bi inter-helix distance for each Cyp 40 TPR is between 7.5 and 8 Å. The closeness of the helices and the narrow omega angle are a consequence of the short 3-amino acid inter helix linker and the presence of the two most strongly conserved residues Ala/Gly8 and Ala20. The key to this tight helix-turn-helix geometry is the stabilising and close-packed interactions between C-beta of Ala-20 and carbonyl oxygen atoms of residues 8, 16 and 17 and the C-β atoms of the (normally) large amino acids 11 and 12 (Figure 31). This sharp turn also brings the side chain of residue 8 close to the backbone of helix B and explains the requirement of the small Gly or Ala side chain at this position. Rather than the classic 'knobs into holes' packing, this is more of a 'staggered small knob' arrangement.

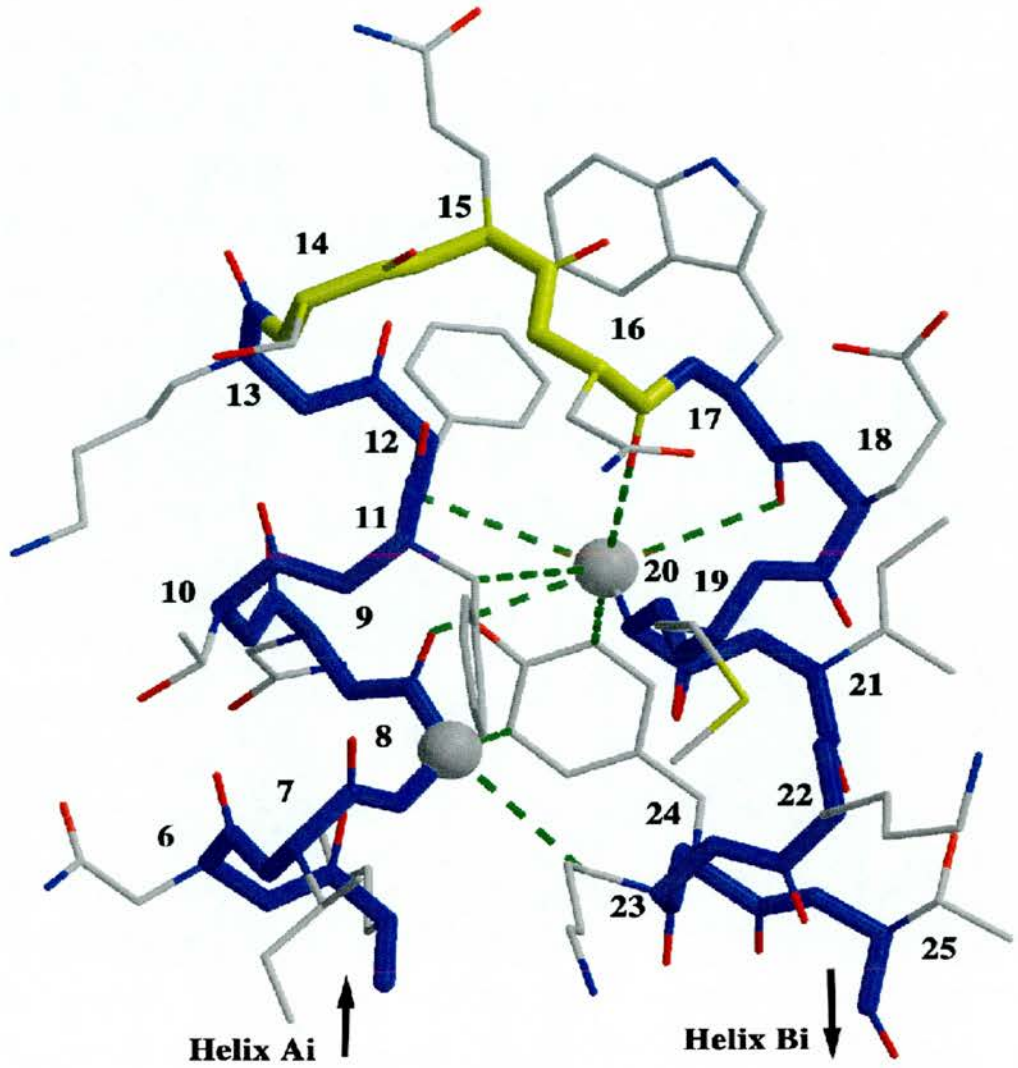


Figure 35

Geometry and intra-TPR motif interactions between helix A_i and B_i.

The geometry of the turn between the A_i and B_i helices is conserved in the three TPR motifs of Cyp 40. A portion of TPR1 from Cyp40 is shown with residues numbered according to the relative TPR-numbers (1 to 34) defined in Figure 31. The tight turn between the helices is governed by the (conserved) conformation of the 3-amino acid inter helix linker (coloured yellow) and the presence of the two most strongly conserved residues Ala/Gly8 (Cα shown as a solid sphere) and Ala20 (Cβ shown as a solid sphere). The tight helix-turn-helix geometry is stabilised by close-packed interactions, shown as dashed green lines, between C-β of Ala-20 and carbonyl oxygen atoms of residues 8, 16 and 17 (with distances between 3.4 and 3.6 Å) and the C-β atoms of the (normally) large amino acids 11 and 12. The close contacts between Gly8 Cα and the side chain atoms of B_1 are also shown.

4.7.4.1

TPR repeat geometry produces a super-helical structure

The inter helix Ω-angles (Bowie, 1997) between adjacent helices are all positive. The interaxial angles between the pairs of helices comprising the three TPR (A_i,B_i) motifs average to 13^0, while the angles between helices of adjacent TPR motifs (B_i, A_{i+1}) average at 24^0. Figure 36 shows a superposition of the structures of Cyp 40 and PP5 (Das et al., 1998) illustrating a very similar arrangement of helices.

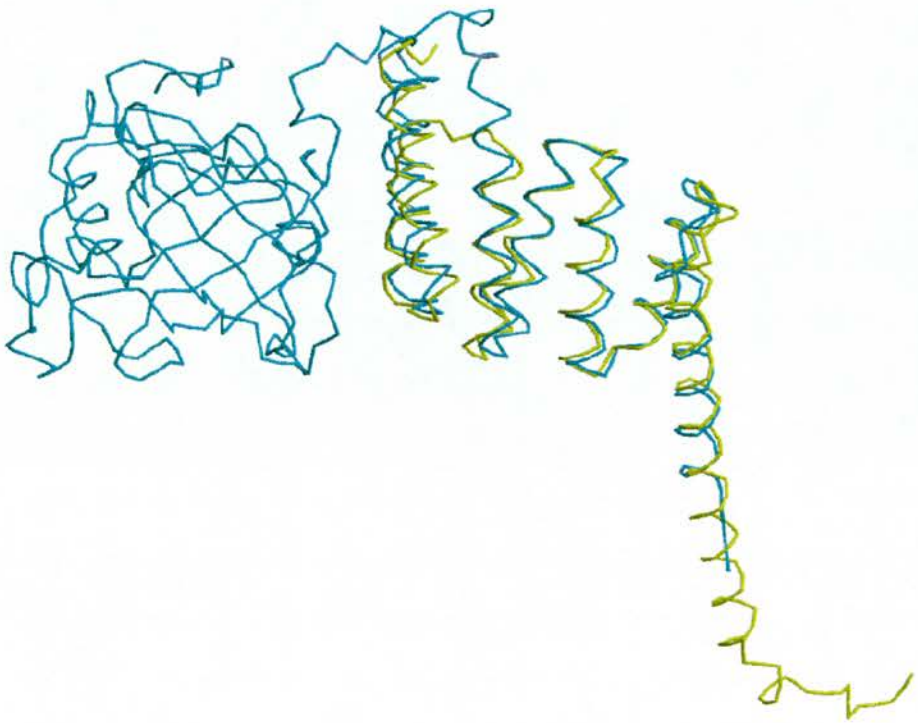


Figure 36

View of the r.m.s.-fit of the X-ray structure of the N-terminal TPR-domain of PP5 and the C-terminal helical domain of Cyp40.

The main chain trace of the X-ray structure of residues 18 to 181 of human Ser/Thr phosphatase PP5 is shown in yellow. Cyp 40 is shown in cyan. A sequence alignment of the TPR motifs for the two structures is given in Figure 31. Those corresponding Cα atoms give an rms-fit of 1.2 Å.

Modelling studies have shown that multiple TPR repeats with alternating Ω angles near 13^0 and 24^0 leads to a right-handed super-helical structure (Das et al., 1998). The two anti-parallel helices of each TPR motif are connected by a 3-residue linker.

In two cases, the 4-residue linkers joining adjacent TPR- motifs adopt type-1 β-turns with proline (Pro 32) at the i+1 position (Wilmot and Thornton, 1988). Differences in inter-helical linker-length helps engineer the different inter-helical angles with the shorter linker constraining a narrower Ω angle between helices A_i and B_I

When sequences of TPR1, TPR2 and TPR3 are compared among different proteins stronger homologies emerge and suggest structural and functional roles for certain amino acids. The nature of the helix packing in tandem repeat structures may require subtly different structural roles for amino acids in the central and flanking TPR motifs. For example there are no inter-helix hydrogen bonds in TPR2, however there are conserved A-B inter-helix hydrogen bonds in TPR1 (Asn231(15)/ND1...Tyr246(24)/OH= 2.64 Å) and there is a conserved salt bridge (Arg313(7)...Asp329(23)) in TPR3. There are a total of 9 inter-helical hydrogen bonds in the TPR domain mostly involving charged residues. A majority of the van der Waals contacts are between the A/B anti-parallel helices of each TPR motif (P/Q make 15 contacts , R/S make 7 and T/U make 17). The close-packed arrangement of helices P, Q and R use the side chains of Y246 and V249 of helix B to fill and pack the centre of the core. Y246 (24) is conserved in all Hsp90 binding proteins from yeast through mammals.

The surface of the Cyp 40 TPR domain shows a pronounced charge distribution (Figure 37 A and B) in which the inner groove of the TPR domain forms a positively charged surface while the outer groove forms a negatively charged surface.

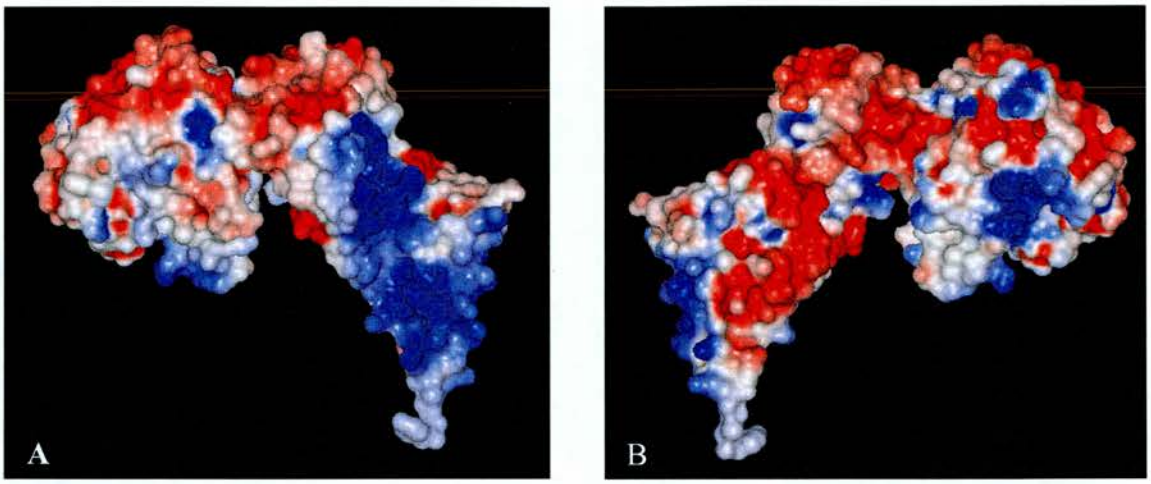


Figure 37 A and B

Representation of the surface electrostatic potential of Cyp 40.

The pronounced charge distribution can clearly be seen, with the inner groove of the TPR domain (A) being positively charged (blue) and the outer groove (B) forming a largely negatively charged surface (red). The figures were prepared using the DelPhi module from the INSIGHT II package.

4.7.5

Helix 'V' may be involved in calmodulin binding

Specific binding between large immunophilins and calmodulin has been shown in bovine Cyp 40 (Ratajczak et al., 1995), maize FKBP 66, (Hueros et al., 1998) and rabbit FKBP52 (Massol et al., 1992). Structural features responsible for calmodulin binding have been determined as: an amphiphilic α-helix that carries a net positive charge (Ghirlanda et al., 1998). The calmodulin binding site to FKBP52 has been proposed to lie between residues 398-415 and corresponds to residues 353-370 of Cyp 40. The X-ray structure of Cyp 40 presented here shows that this putative interaction domain corresponds to the C-terminal portion of the V-helix (Figure 32) which is seen to extend beyond the shorter TPR helices. The protruding and predominantly positively charged helix-V would therefore be accessible for calmodulin binding.

4.8

STRUCTURE/FUNCTION INSIGHTS FROM CYP 40 TPR DOMAIN GEOMETRY

4.8.1

Structural underpinning of the TPR motif

The X-ray structures of TPR proteins characterised to date provide a basis for understanding the pattern of residue conservation seen. A consensus sequence pattern of small and large hydrophobic residues was postulated with small residues occurring at positions 8, 20 and 27 (Das et al., 1998). The influence of two of these conserved residues (at positions 8 and 20 in the TPR motif sequence) on helix packing was discussed in section 4.7.4, explaining the requirement for these small residues in engineering the geometry of the tight turn between the A_i and B_i helices.

A more elaborate sequence of loosely conserved residues was originally suggested (Sikorski et al., 1990). These authors proposed that when comparisons of size and nature were made between amino acids, a largely conserved pattern of similarities could be determined from degenerate sequences. In particular they suggested eight of the 34 amino acids were more highly conserved, generating a consensus sequence for positions 4 (W,L,F), 7 (L,,I,M), 8 (G,A,S), 11 (Y,L,F), 20 (A,S,E), 24 (F,Y,L), 27 (A,S,L) and 32 (P,K,E).

Possible functional specialisation within the TPR motif was suggested in this paper, an idea further developed in a short review of TPR containing proteins (Blatch and Lassle, 1999).

4.8.2

The TPR motif of Cyp 40 is most divergent amongst those structurally characterised

4.8.2.1

Comparisons between structures

Analysis of TPR domain structures (presented in Table 12) and sequence alignments (illustrated in Figure 38), reveal Cyp 40 diverges most in overall structure from the classical helix (13aa) turn (3aa) helix (14aa) TPR motif.

The first 3 helices of the Cyp 40 TPR domain are significantly elongated when compared with those from other structurally characterised TPR containing proteins. Cyp 40 shares with PP5 and TOM 20 an elongated A helix in TPR 1, the B helix of TPR 1 is also elongated by one turn with respect to the remaining proteins. The A helix of TPR 2 in Cyp 40 is also significantly elongated, no other structurally characterised protein has that feature.

The remaining units of the motif comply fairly well with the classical definition of helix turn helix motif seen in the structure of PP5 (Das et al., 1998).

Protein	Structure paper reference	PDB ID NO.	Binding partners /function	No. TPR repeats	Location of TPR repeats in sequence (residue no.)	Length of helices in structure (residues)
Cyp 40 (monoclinic) CYP4_BOVIN	(Taylor et al., 2001)	1IHG	Cyclophilin /PPIase Hsp90/SHR's	3	223 – 256, 273 – 306 308 – 340	P(20), Q (21), R(24), S(12), T(13), U(14), V(18)
Cyp 40 (tetragonal) CYP4_BOVIN	(Taylor et al., 2001)	1IIP	Cyclophilin /PPIase Hsp90/SHR's	3 (2)	223 – 256, 273 – 306 308 – 340	P(20), Q (21), R(24), S(12)
PP5 PPP5_HUMAN	(Das et al., 1998)	1A17	Ser/Thr phosphatase Hsp90	3	28 – 61, 62 – 95 96 – 129	A1(19), B1(14), A2(13), B2(14), A3(13), B3(14)
p67 Phox NCF2_HUMAN	(Lapouge et al., 2000)	1E96	Subunit of NADPH oxidase. Binds Rac.	4	37 – 70, 71 – 104, 121 – 154	A1(16), B1(9), A2(13), B2(14), A3(13), B3(13), A4(14), B4(13)
Hop/TPR2 Hsp90 **Hop/TPR1** Hsp70 IEFS_HUMAN	(Scheufler et al., 2000)	1ELR 1ELW	**H**sp 90 **O**rganiser **P**rotein	9	4 – 37, 39 – 71, 73 – 105, 225 – 258, 260 – 292, 300- 333, 360 – 393, 395 – 427, 428 – 461	A1(15), B1(14), A2(14), B2(17), A3(17), B3(14) A1(14), B1(14), A2(14), B2(14) A3(13), B3(12)
PEX5 **PEX5** (extended) PEX5_HUMAM	(Gatto et al., 2000) (Kumar et al., 2001)	1FCH 1HXI	Receptor for the peroxisomal targeting signal PTS-1	7 (6) 7 (3)	299 – 331, 332 – 365, 366 – 399, 415 – 448, 451 – 484, 485 – 518, 519 - 552	A1(10), B1(13), A2(13), B2(14) A3(14), B3(14) A1(10), B1(14), A2(13), B2(14) A/B3(28)
TOM20 OM20_RAT	(Abe et al., 2000)	1OM2	Receptor subunit of **T**ranslocase of the **O**uter **M**itochondrial membrane	1	63 - 103	A1 (21), B1 (15)

Table 12 (above)

Analysis of structurally characterised TPR domains.

Table 12 summarises data available on all structurally characterised TPR containing proteins. A search of the Research Collaboratory for Structural Bioinformatics (RSSB) PDB database for "tetratricopeptide" containing proteins identified six of the above structures. A further search for "TPR" identified 24 structures, including the above, however closer examination of the database entries additionally identified Hop only as containing a recognisable TPR domain. Vesicular transport protein Sec 17 was also listed, due to local structural similarity to TPR repeat proteins. The Sec 17 structure however has a different overall twist (Rice and Brunger, 1999). Lengths of helices were determined using data from the PDBSum database. This database provides a summary of all the secondary structure elements contained within a protein and hence is a useful tool in the analysis of a number of structures.

```
CYP4_BOVIN   (216-259)223-256      VDKILLI SEDLKNIGNT   FFKSQNWEMA   IKKYTKVLRY   VEGSRAA
PPP5_HUMAN    28- 61            PPADGALKR AEELKTQAND   YFKAKDYENA   IKPYSQAIEL   NPSN
IEFS_HUMAN   225-258                      ALKEKELGND   AYKKKDFDTA   LKHYDKAKEL   DPTN
NCF2_HUMAN     3  36                      LVEAISLWNE   GVLAADKKDW   KGALDAFSAV   QDPH
NCF2_HUMAN    37- 70                      SRICFNIGCM   YTILKNMTEA   EKAFTRSINR   DKHL
PEX5_HUMAN   299-331                      HPQPFEEGLR   RLQEGDLPNA   VLLFEAAVQQ   DPKH
OM20_RAT      70-103              EAVQKF FLEEIQLGEE   LLAQGDYEKG   VDHLTNAIAV   CGQP
TPR 2 OF SELECTED PROTEINS
CYP4_BOVIN   (262-285)273-306 DADGAKLQPVA LSCVLNIGAC   KLKMSDWQGA   VDSCLEALEI   DPSN
PPP5_HUMAN    62- 95                      AIYYGNRSLA   YLFTECYGYA   LGDATRAIEL   DKKY
IEFS_HUMAN   300-333                      MTYITNQAAV   YFEKGDYNKC   RELCEKAIEV   GREN
NCF2_HUMAN    71-104                     AVAYFQRGML   YYQTEKYDLA   IKDLKEALIQ   LRGN
PEX5_HUMAN   332-365                      MEAWQYLGTT   QAENEQELLA   ISALRRCLEL   KPDN
TPR 3 OF SELECTED PROTEINS
CYP4_BOVIN   307-340                      TKALYRRAQG   WQGLKEYDQA   LADLKKAQEI   APED
PPP5_HUMAN    96-129                     IEGYYRRAAS   NMALGKFRAA   LRDYETVVKV   KPHD
IEFS_HUMAN   259-292                      AKAYARIGNS   YFKEEKYKDA   IHFYNKSLAE   HRTP
NCF2_HUMAN   121-154                     ACEVLYNIAFM   YAKKEEWKKA   EEQLALATSM   KSEP
PEX5_HUMAN   366-399                      QTALMALAVS   FTNESLQRQA   CETLRDWLRY   TPAY
TPRS 4-7 OF SELECTED PROTEINS
IEFS_HUMAN   225-258                      ALKEKELGND   AYKKKDFDTA   LKHYDKAKEL   DPTN
PEX5_HUMAN   415-448                      LGPSKRILGS   LLSDSLFLEV   KELFLAAVRL   DPTS
IEFS_HUMAN   260-292                      TYITNQAAVY   FEKGDYNKCR   ELCEKAIEVG   REN
PEX5_HUMAN   451-484                      PVDQCGLGVL   FNLSGEYDKA   VDCFTAALSV   RPND
IEFS_HUMAN   300-333                      AKAYARIGNS   YFKEEKYKDA   IHFYNKSLAE   HRTP
PEX5_HUMAN   485-518                      YLLWNKLGAT   LANGNQSEEA   VAAYRRALEL   QPGY
IEFS_HUMAN   360-393                      ALEELNKGNE   CFQKGDYPQA   MKHYTEAIKR   NPKD
PEX5_HUMAN   519-552                      IRSRYNLGIS   CINLGAIREA   VEIIFLEALNM  QRKS
IEFS_HUMAN   395-427                      AKLYSNRAAC   YTKLLEFQLA   LKDCEECIQL   EPTF
IEFS_HUMAN   428-461                      IKGYTRKAAA   LEAMKDYTKA   MDVYQKALDL   DSSC
```

Figure 38

Sequence alignment of TPR domains from structurally characterised proteins, illustrating the differences seen between proteins in helix length. Proteins were extracted from the SWISS-PROT database and manually aligned over the TPR domain. The proteins illustrated are bovine Cyp 40 (CYP4_BOVIN), PP5 (PPP5_HUMAN), Hop (IEFS_HUMAN), $P67^{phox}$ (NCF2_HUMAN), PEX5 (PEX5_HUMAN) and TOM20 (OM20_RAT).

4.8.3

Analysis of X-ray and sequence data allows dissection of structurally and functionally significant TPR motif residues

The availability of six high resolution X-ray structures of TPR containing proteins, combined with sequence data on the Hsp 90 binding TPR's; allowed some dissection of the contribution of individual "conserved" residues to maintenance of structural integrity of the molecule whilst also possibly providing insights into the nature of the interaction with Hsp90.

4.8.3.1

Conservation of specific residues within a "functional subfamily" of TPR proteins

Position	TPR 1	TPR 2	TPR 3
2.			***K (9/9)*** K (3/5)
3.			A (3/5)
5.	***K (9/9)***		**Y (7/9)** Y (3/5)
6.		***N (9/9)***	***R (9/9)*** R (3/5)
7.			**R (7/9)**
8.	**G (8/9)**		A (4/5)
10.			S (3/5)
11.		Y (3/5)	
12.	**F (7/9)**	**L (7/9)**	
13.	***K (9/9)***		
20.	***A (9/9)*** A (5/7)	A (4/5)	***A (9/9)*** A (5/5)
23.			**D (7/9)**
24.	***Y (9/9)***		
26.	**K (7/9)**		
27.		**A (7/9)** L (4/5)	
28.		**L (7/9)** L (3/5)	
29.		E (4/5)	
32.			**P (7/9)** P (3/5)
34.		N (4/5)	

Table 13

Conservation of specific residues within Hsp90 binding proteins (see Figure 31). Residues completely conserved between proteins in Figure 31 are bold, italicised and underlined, largely conserved residues are illustrated in bold. Residues conserved between structurally characterised proteins are also illustrated on the right hand side of each column. Numbering refers to the how many representatives of each of the residues are conserved for a given position number (ie at position 5 in TPR 1 from all the protein sequences illustrated in Figure 31, K is completely conserved hence K (9/9).

Analysis of the sequences including alignment into individual TPR repeats (see Table 13) reveals a higher degree of conservation of residues within a functionally specific family i.e. the Hsp90 binders compared to those structurally characterised.

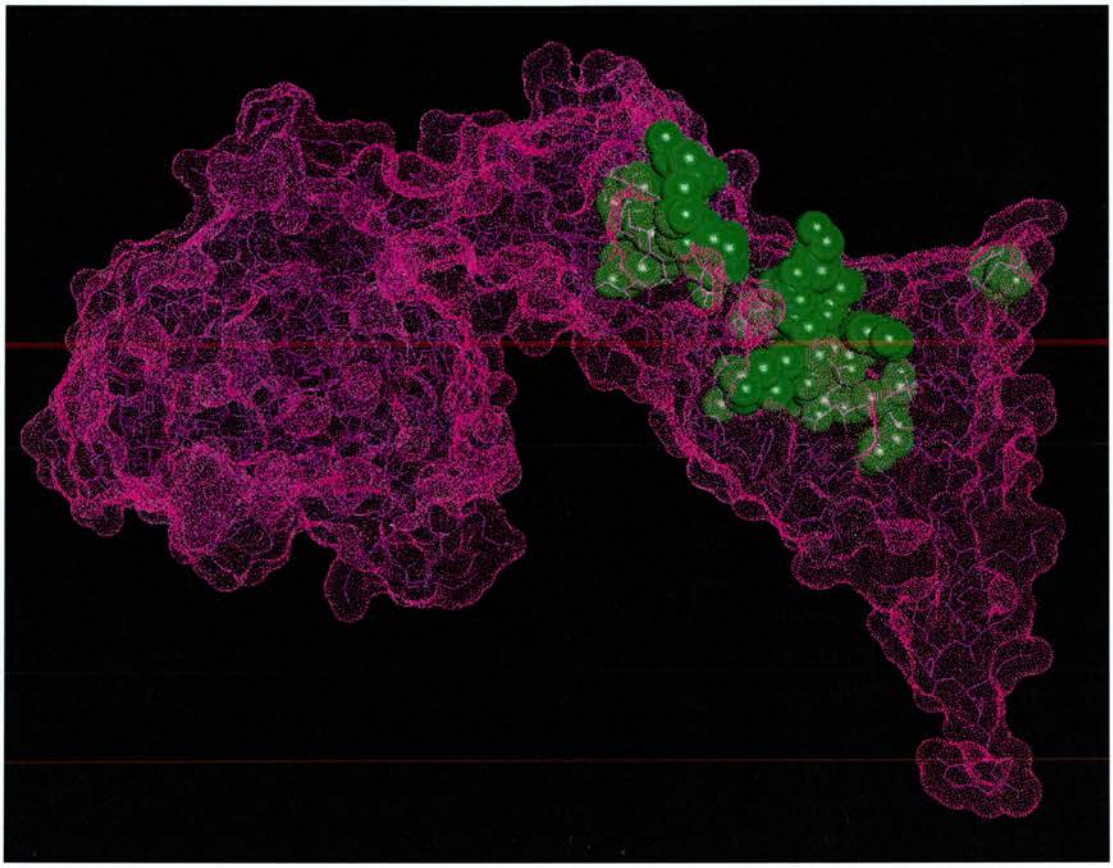


Figure 39

Residues conserved between members of the Hsp90 binding sub-family (see Table 13) were highlighted in the Cyp 40 sequence and displayed with space filling CPK. Conserved residues cluster in the binding groove, providing further evidence in support of this region being important in interactions between these two proteins. The majority of the residues are located within the positively charged inner groove, 2 residues being within the outer negatively charged groove and 1 (Pro 338) between last 2 helices. The figure was prepared using WitnotP (Widmer, 1997).

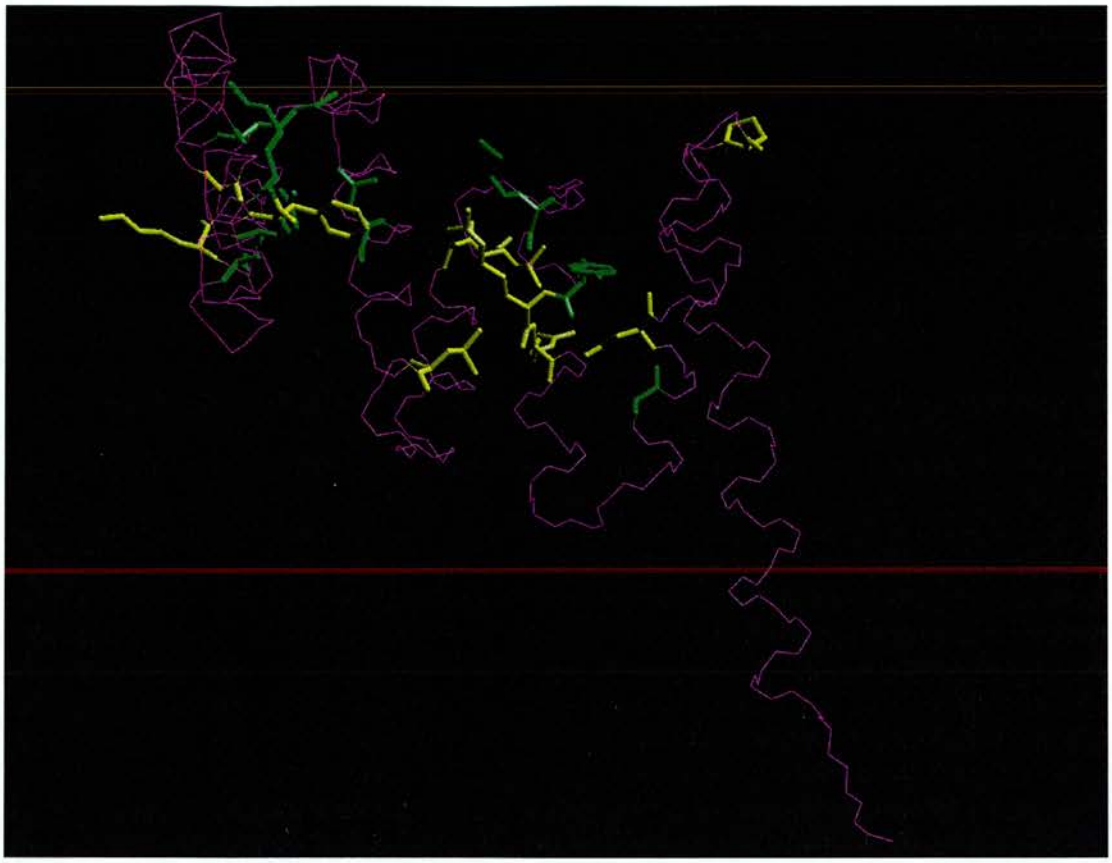


Figure 40

Cyp 40 TPR displayed with backbone atoms only. Conserved residues from Table 13 and Figure 38 above displayed with side chains on. Completely conserved residues are coloured green, largely conserved residues are coloured yellow. The majority of the residues point into the binding groove and consequently would be available for interactions with partner proteins. A number of residues, particularly Ala and Gly appear to be more intimately associated with conservation of structural integrity of the TPR domain, as discussed earlier.

Amongst the structurally characterised proteins, TPR 3 is most similar, this being largely weighted by the effect of the Hsp 90 binding proteins predominating in the alignments These results are suggestive of functional requirement for these particular amino acids in the interaction with the whole Hsp90 molecule. Mutagenesis studies might help to identify if the residues highlighted above are indeed critical for the interaction between these two proteins.

4.9

DISCUSSION

The two very different structures of Cyp 40 provide unique and potentially extremely interesting insights into the folding pathway of TPR containing proteins and/or an active functional conformation of Cyp 40. One hypothesis is that of the tetragonal form being a folding intermediate, trapped on the way to the fully folded conformation. An interesting possibility, for which there is as yet no evidence, could involve the extended conformation also being an active conformation. Large scale movement of the TPR domain, as seen between the two structures, may possibly be triggered by substrate binding and release. A recent paper reporting the crystal structure of a highly helical, flexible, chaperone protein Prefoldin (Siegert et al., 2000) raises the intriguing possibility of the extended conformation being an " active chaperoning" structure. However, as Cyp 40 seems to be largely monomeric in solution, it seems unlikely that extended Cyp 40 would function alone. Partner proteins from a hypothetical active chaperone complex remain to be characterised. Further experiments involving formation of a complex with substrate proteins may shed further light onto this possibility.

The high resolution X-ray structures of two crystal forms of bovine Cyp 40 provide further insights into the versatile nature of the TPR repeat motif. Whilst the overall super-helical structure is maintained by these proteins, diversity can be seen between structures. Results from comparisons between Cyp 40 and other TPR containing proteins are indicative of this motif having greater structural flexibility than originally envisaged. The TPR motif in Cyp 40 contains the longest helices yet determined. The functional significance of these elongated helices remains to be established. One possible explanation might include "fine tuning" of a binding groove by engineering the interaction surface of the TPR domain by slight alterations in the spatial arrangements of each TPR motif. An example of such specificity of interaction can be seen in the hierarchy of binding affinities seen with Hsp 90 and members of the large immunophilin family (Carrello et al., 1999).

Another recent paper reported an "unexpected extended" conformation for the third TPR motif of the Peroxin PEX5 from *Trypanosoma brucei*, in which the last two helices are folded into one single 44Å long continuous helix (Kumar et al., 2001). This structure differed from the first structure of Pex5, suggestive of different "open" and "closed" conformations of the protein. Possible roles for these "open" and "closed" conformations of the TPR motif were suggested, with a role for the extended conformation in interactions with other peroxins or whilst undergoing membrane translocation.

An overlay of the three structures (Cyp 40 monoclinic, tetragonal and PEX5 extended) showed the PEX5 to have a very similar conformation to that of monoclinic Cyp 40. The length of the C-terminal helices of monoclinic Cyp 40 and PEX5 were very similar at 42Å and 44Å respectively, although the Cyp 40 C-terminal helix adopts a more angular structure.

Characterisation of bovine Cyp 40 chaperone activity

Preliminary results from a project investigating chaperone activity in cyclophilins distinct from PPIase activity, (carried out as part of a Carnegie summer studentship held by Simon Wilkinson) provided evidence in support of Cyp 40 acting as a molecular chaperone. To analyse the contribution from the PPIase domain alone, this study utilised Cyp 3 (as discussed in Chapter 3) as a representative of a single domain cyclophilin, compared and contrasted with Cyp 40. Incubation with Cyp 40 but not Cyp 3 was shown to slow the rate of aggregation of thermally denatured citrate synthase. Complete suppression of aggregation was seen when GroEL was used.

The chaperone activity of Cyp 40 was first characterised in 1996, maintaining an unfolded substrate in a folding competent intermediate, lacking the ability to completely refold substrate (Freeman et al., 1996). Data from the above experiments do not produce evidence for any additional roles for Cyp 40 in chaperone interactions.

Subsets of the Cyp 40 protein family

Specialisation of functions of Cyp 40 type proteins has been suggested from studies comparing activities of Cpr6 and Cpr7, two Cyp 40 homologues present in *Saccharomyces cerevisiae*. Following heat shock, expression patterns were different for these two proteins. Their relative PPIase activities have also been studied (Warth et al., 1997; Dolinski et al., 1997; Duina et al., 1998b). An analysis of the functional properties of these two proteins found Cpr6 and Cpr7 differed in their stability to thermal denaturation, their ability to isomerase Xaa-Pro peptide bonds and in their molecular chaperone activities (Mayr et al., 2000). The PPIase activity of Cpr7 was about 6-fold lower than that of Cpr6, however Cpr7 was a more potent chaperone, suggestive of significant functional specialisation between the two proteins.

Bovine Cyp 40 more closely resembles Cpr6 than Cpr7, sharing 44% identity over the whole protein (58% identity over the cyclophilin domain and 31% identity over the TPR domain). The limited degree of chaperone activity seen with Cyp 40 may be reflective of the subset of Cyp 40 type proteins which has evolved over time for a different regulatory function.

CHAPTER 5.

PURIFICATION AND CHARACTERISATION OF C-TERMINAL FRAGMENTS OF HEAT SHOCK PROTEIN 90

5.1

INTRODUCTION

Heat shock protein 90 is a ubiquitous, highly conserved, key regulatory molecule involved in many important processes including signal transduction events, control of elements of the cell cycle and regulation of transcription.

5.1.1

The Hsp 90 binding immunophilins

Isolation of Hsp90 from a preparation of *S.pombe* cell extract through a bioaffinity interaction with Cyp 40 was described in Section 4.5.2, Chapter 4. The interaction between large immunophilins and Hsp 90 had been previously described (Callebaut et al., 1992; Lebeau et al., 1992; Tai et al., 1992; Ratajczak et al., 1993).

This chapter will introduce the Hsp90 protein family, summarise what is known to date of their structure and functions and conclude with sections focussing on the isolation and characterisation of C-terminal fragments of Hsp90 β. The current lack of structural data on the interactions of Hsp90 with partner proteins, suggested molecular modelling of possible interactions might provide insight into the problem. Modelled interactions between Cyp 40 (using the Cyp 40 structure described in Chapter 4 as the basis) and Hsp90 will be considered in greater detail in Chapter 6

5.1.2

Identification of the Hsp90 family

A number of proteins were initially identified by their marked up-regulation in cellular concentration following heat stress. Amongst those identified were members of what would become known as the Heat Shock Protein 90 family (Hsp90) (Borkovich et al., 1989). This family of proteins is highly abundant in the cell, constituting approximately 1-2 % of soluble cellular protein under normal conditions and is represented by numerous isoforms, (Welch and Feramisco, 1982; Gupta, 1995; Lindquist and Craig, 1988).

Two genes encode for different, closely related mammalian isoforms of Hsp90 (α and β), with the recently isolated mitochondrial homologue TRAP1/Hsp75 (Felts et al., 2000) being less well conserved and also less abundant (Song et al., 1995) and (Chen et al., 1996a). Homologues of Hsp90 can also be found in the endoplasmic reticulum (Argon and Simen, 1999; Stevens and Argon, 1999) and chloroplast (Schmitz et al., 1996). The degree of conservation between homologues is high, Hsp90α being 86% homologous to Hsp90β (Hickey et al., 1989). A high degree of conservation is also maintained over a variety of species, human Hsp90β shares 78% homology with its homologue from *D. melanogaster* and is 61% homologous with yeast Hsp90 (Rebbe et al., 1987). The *E.coli* homologue HtpG is approximately 42% homologous to Hsp90β, lacking the charged linker region.

Hsp 90 has been shown to be essential for viability in *D. melanogaster* (Cutforth and Rubin, 1994; Yue et al., 1999) and in the yeast *S. cerevisiae* (Aligue et al., 1994; Borkovich et al., 1989). It has been suggested that for efficient protein folding in the endoplasmic reticulum, GRP94 (the ER resident Hsp90) is essential (Gidalevitz et al., 2001).

A huge amount of research effort has focused on the interaction mechanisms of this important molecule and its "client" or partner proteins.

5.1.3

Domain structure of Hsp90 from sequence and biophysical data

Sequence alignment studies on members of the Hsp90 family revealed a high degree of conservation between family members. Hsp90 has a modular domain structure in which the highly conserved N and C-terminal domains, approximately 25 and 55kDa respectively, are joined by a highly charged linker. This linker region is however, non-essential, and is absent in both the *E. coli* Hsp90 homologue HtpG (Spence and Georgopoulos, 1989) and eukaryotic mitochondrial homologues (Felts et al., 2000).

Native Hsp90 is a large molecule and continues to represent a considerable challenge for structural biologists. Early attempts to structurally characterise Hsp90 included a number of biophysical and biochemical studies on the whole molecule. Recently X-ray crystallographic studies have focused on domains of Hsp90, with some degree of success (for a review of Hsp90 structural data see (Pearl and Prodromou, 2000). The results of some of these experiments will be briefly summarised below.

Mammalian Hsp90 was used in electron microscopic studies of the protein and showed native Hsp90 was present as an elongated dimer. Dimerisation occurred through the extreme carboxy terminal domain. This suggested Hsp90 was a very flexible molecule as a number of reports had observed dimerisation between the N-terminal domains.

Labelling the Hsp90 molecule with monoclonal antibodies raised against N and C-terminal regions provided further insights into its structural organisation. The model proposed from these studies was of two Hsp90 monomers aligning in an anti-parallel orientation, with dimerisation occurring through the C-terminal domain, leaving the N-termini of the dimer free (Maruya et al., 1999). This group also reported changes in morphology of the Hsp90 homo-dimer into a ring like structure following addition of ATP or exposure to heat stress, once again suggestive of great structural flexibility within the molecule.

The endoplasmic reticulum homologue GRP94 has also been studied, again dimerisation was shown to be mediated *via* the C-terminal domain (Wearsch et al., 1998).

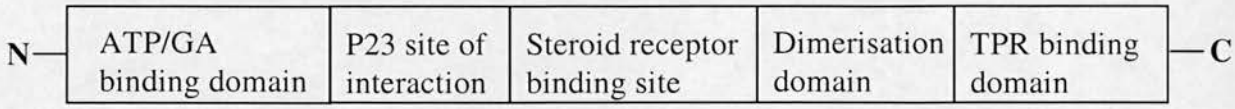


Figure 41

Cartoon representation of the five functional domains in the structure of Hsp90 (not drawn to scale). The domains are indicated separately in the figure, however significant functional overlap exists between domains.

5.1.4

The C-terminal domain of Hsp90 interacts with TPR containing proteins

The C-terminal region of Hsp90 is critical for interaction with TPR motif containing cofactors. Using the yeast two hybrid system, the TPR interaction region in Hsp90 was delineated to its C-terminal 12kDa domain, a result subsequently confirmed by bioaffinity interaction binding assays (Chen et al., 1996b; Young et al., 1998). Different members of the TPR protein family associate with Hsp90 in a sequential manner, the order and specificity of the interaction being critical for proper protein folding (Frydman and Hohfeld, 1997; Johnson and Craig, 1997). Specificity of Hsp90 binding was investigated for three different TPR containing proteins; PP5, FKBP52 and Hop. Mutation of conserved residues from the extreme C-terminus of Hsp90 (Glu729, Glu730 and Asp732) inhibited binding of all three proteins, mutation of other residues had more specific effects (Ramsey et al., 2000). This overlapping interaction domain in Hsp90 for immunophilins, PP5 and Hop suggested a common TPR acceptor site in Hsp90 (Russell et al., 1999).

The binding domain in Hsp90 for Cyp 40 and Hop was also identified from a yeast two hybrid screen. A mouse cDNA library was screened, all interacting clones contained the intact carboxyl terminus of Hsp90, overlapping a common region corresponding to amino acids 558-724 of murine Hsp90. Further studies delineated the critical interaction region to lie within a 124 residue carboxy terminal segment of Hsp 90. Deletion mutants in which the highly conserved sequence at the extreme carboxyl terminus of Hsp90 (EEVD) was removed precluded interaction with both Cyp 40 and Hop, underlining the importance of this motif in the interaction (Carrello et al., 1998).

5.2
X-RAY STRUCTURE OF THE N-TERMINAL DOMAIN OF HSP 90

Recently structures of N-terminal fragments of both human and yeast Hsp90 have been solved by X-ray crystallography.

5.2.1
Human Hsp90

Human Hsp90, residues 9-236, was first crystallised as a complex with the anti-tumour antibiotic geldanamycin (GA). The interaction between Hsp90 and geldanamycin had been characterised (Whitesell et al., 1994). Studies investigating GA inhibition of certain cell cycle kinases showed GA binding to Hsp90 only. Resultant indirect inhibition of kinase activity was due to failure of Hsp90 dependent folding, (Whitesell and Cook, 1996). The precise nature of the interaction between GA and Hsp90 was however, unknown. To further characterise this interaction; the 3 domains of Hsp90 were assayed separately for their GA binding activities, the N-terminal domain was shown to bind GA at an approximately equimolar ratio (Stebbins et al., 1997).

Crystallisation trials resulted in two crystal forms. "Free" Hsp90 produced two different crystal forms, the Hsp90/GA complex crystallised only in one of the aforementioned forms (Stebbins et al., 1997). The site of GA binding was shown to be in an approximately 15Å deep pocket. Residues clustering in and around this binding pocket are highly conserved over diverse species, 82% of the residues lining the binding pocket interior are conserved from *E. coli* to humans. The authors went on to speculate the structures obtained from the two different crystal forms represented the binding pocket trapped in "open" and "closed" conformations, with consequent implications on accessibility for ligand binding.

In the crystal structure, GA was bound in a very compact conformation, leading to the suggestion that the conformation of bound ligand mimicked that of a potential natural substrate, a polypeptide from a "client protein". Another important feature identified by these authors was the conservation of residues that form a surface groove and may be important in Hsp90 protein-protein interactions.

5.2.2

Yeast (Saccharomyces cerevisiae) Hsp90

The crystal structure of residues 1-220 of *S. cerevisiae* Hsp90 (Hsp82) was solved in two crystal forms. Hsp90 was present as a dimer in both crystal forms. Dimerisation occurred between the C-terminal β-strands of each monomer, with further stabilisation of the dimer interface by non-covalent interactions (Prodromou et al., 1997a). The two crystal forms grown once again captured the molecule in both "open" and "closed" conformations, illustrative of the Hsp90 dimer having the potential to act as a "molecular clamp", binding fragments of a polypeptide chain or peptides. The need for great structural flexibility of the molecule was also discussed in this paper. It was suggested that interactions occurring with other proteins in different regions of the intact Hsp90 molecule would lead to conformational changes in the molecule. The physiological relevance of the structure described by Prodromou et al is yet to be proven, indeed in a follow-up paper from the same group the validity of the potential peptide binding channel defined from the dimer structure

was questioned, as the authors were unable to detect binding of peptides to this channel.

The structure of the N-terminal domain of both human and yeast Hsp90 was shown to be remarkably similar, as may have been expected from the high degree of conservation seen between the sequences. Crystallographically, human protein is monomeric, yeast dimeric. The overall structure of each molecule however is that of an α /β sandwich, consisting of eight stranded anti-parallel β sheet forming a potential binding groove, with 9 α helices on one face of the molecule, with the sides of a deep binding pocket formed by the nine α helices.

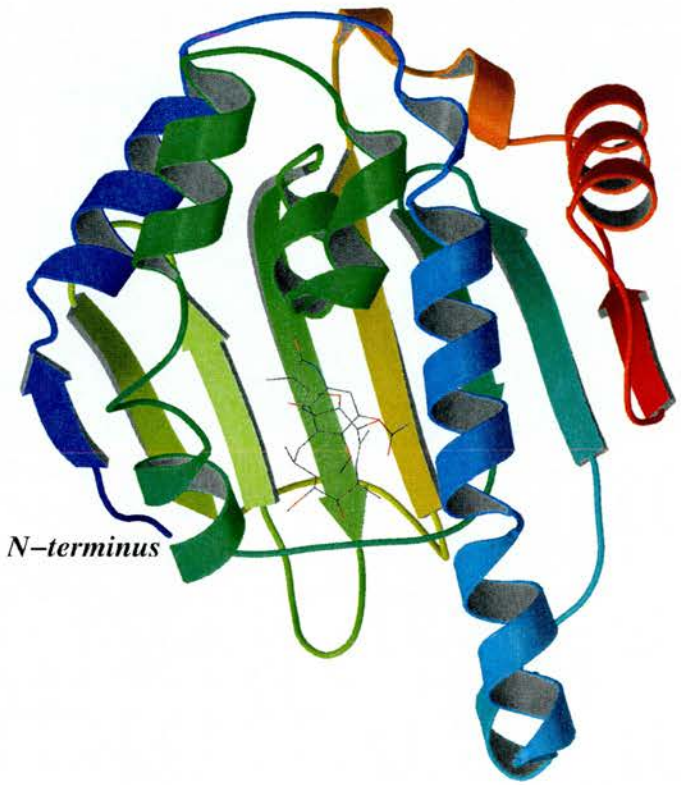


Figure 42

Structure of the N-terminal domain (residues 9-236) of human Hsp90 in the open conformation; with geldanamycin bound deep in the binding pocket (Stebbins et al., 1997). The figure was prepared using PDB ID 1YET as the template and was rendered using Molscript (Kraulis, 1991).

5.2.3

Hsp 90 and ATP

A number of studies showed ATP was necessary for many important processes within the Hsp90 system, including the release of Hsp 90 from substrate complexes (Schneider et al., 1996, Grenert et al., 1997). ATP binding and hydrolysis was shown to be essential for Hsp90 to function *in vivo* and *in vitro*, from both *S. cerevisiae* (Hsp82), (Obermann et al., 1998) and *E. coli* (HtpG) (Panaretou et al., 1998).

The presence of ATP binding to Hsp90 was shown unambiguously by publication of a high resolution crystal structure of a complex between ATP/ADP and the N-terminal domain of Hsp90 (Prodromou et al., 1997b). The ATP binding site identified more closely resembles that seen in DNA gyrase B than a conventional Walker site (Panaretou et al., 1998).

This group went on to report regulation of Hsp90 ATPase activity by TPR containing co-chaperones (Prodromou et al., 1999) and suggested ATP hydrolysis triggered client protein release.

A later paper suggested the ATPase cycle of Hsp90 drives a molecular clamp via transient dimerisation of the N-terminal domains of Hsp90. This situation echoes that seen with both MutL and DNA gyrase. In DNA gyrase the clamp mechanism controls binding and release of a DNA duplex molecule during the topoisomerase strand passage reaction (Kampranis et al., 1999), whilst with MutL the exact mechanism is as yet unclear (Ban et al., 1999).

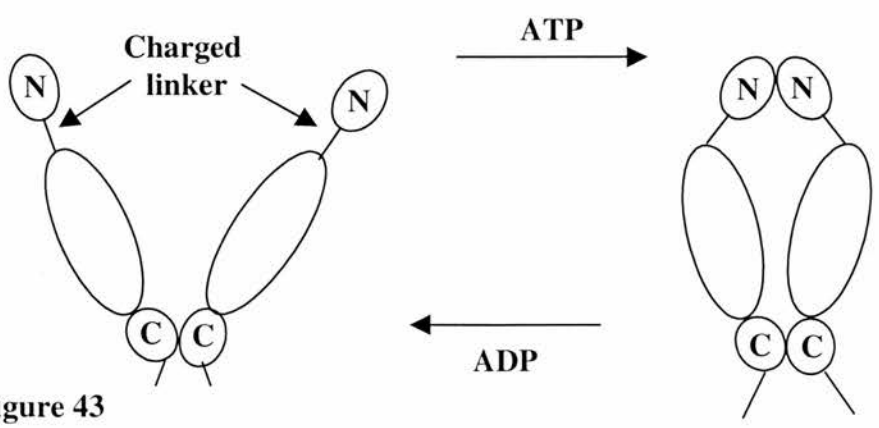


Figure 43

The proposed "molecular clamp" of Hsp90. Opening and closing is coupled to the ATPase cycle.

5.3
FUNCTIONS OF HSP 90

The numerous roles of Hsp90 are reviewed in a number publications including a study of plant Hsp90 functions (Pratt et al., 2001), an overview of the roles of chaperone proteins in the maturation of steroid receptors (Kimmins and MacRae, 2000) and a review of the field of chaperones in mitogenic signal transduction and cell cycle regulation (Helmbrecht et al., 2000). A number of other general reviews are also available including (Caplan, 1999, Buchner, 1999; Pratt, 1998).

5.3.1
Hsp90 associated with protein kinases, cyclins and cyclin dependent kinases

Hsp90 has been shown to be necessary for the proper maturation, intracellular transport and regulation of activity of many protein kinases involved in signal transduction, translation control and haem-dependent nitric oxide synthase. Amongst those that have been shown to require Hsp90 are the Src family protein kinases, the Wee1 tyrosine kinase, serine/threonine kinases Raf-1 and Mek, the cyclin dependent serine/threonine kinase Cdk4 and the haem-dependent eIF-2α kinase (Pratt and Toft, 1997). Hsp90 is often targeted to its kinase substrate by Cdc37/p50 (Hartson et al., 2000), in a situation which mirrors that seen with steroid hormone receptors, Hsp90/kinase complexes act to maintain the kinase in an activation-competent state.
The best-characterised Hsp90 dependent kinases are those of the Src family. The interactions between the Src family and Hsp90 are however very specific, as can be seen from studies on both the viral v-Src and cellular c-Src proteins. Both proteins require Hsp90 for the maturation of their catalytic activity, c-Src also requires Hsp90 in order to function as a substrate for the regulatory kinase Csk.

5.3.2

Transcription factors including p53

Wild type p53 transcription factor is a nuclear tumour suppresor involved in regulation of the cell cycle. Mutations of the p53 tumour suppressor gene are amongst the most common molecular genetics defect found in human cancers. The result of most p53 mutations is to produce misfolded proteins lacking their usual cell cycle checkpoint activity. This results in tumours with high levels of aberrant protein, due in part to the prolongation of the half-life of the misfolded p53 due to impairment in the ubiquitin-proteasome pathway. Sequestration of mutant p53 within the "machinery" of the molecular chaperone complex, with consequent protection from degradation, leading to its eventual accumulation in tumours was suggested (Whitesell et al., 1997).

A role for Hsp90 in the proper folding of mutated p53 was demonstrated, the effect of geldanamycin was also investigated. These studies revealed the importance of functional Hsp90 as addition of GA to reaction mixtures reduced the proportion of mutated p53 translated *in vitro*. It was suggested that Hsp90 was required for the folding and /or stabilisation of the mutated form of p53. Results from both *in vitro* and *in vivo* experiments appeared to confirm this theory (Blagosklonny et al., 1996).

A recent paper (Whitesell et al., 1998) studied the effect of geldanamycin on interactions between mutant p53 and components of the chaperone machinery. A number of proteins co-precipitated in a molybdate sensitive fashion with mutant p53, including Hsp90, Cyp 40 and p23. To establish the role(s) played by Hsp90 in association with p53, GA was once again used to disrupt the interaction. A cell line with a temperature sensitive mutation of p53 was used in this study. At 32°C this cell line produced wild type protein, at 39°C it produced protein with altered conformation, localisation and function. Wild type protein did not, under these circumstances, stably associate with the group of molecular chaperones found in stable association with mutant p53.

5.3.3

Potential role in evolution as 'morphological capacitor'

An interesting series of experiments investigated the phenotypic variation seen in populations of *Drosophila melanogaster* carrying mutations in their Hsp90 genes (*Hsp83*).

Homozygous *Hsp83* mutations are lethal, however a number of groups had observed an increased prevalence of morphological abnormalities in fly stocks maintained heterozygous for the *Hsp83* mutation. Crossing these mutant flies with standard laboratory strains also produced a significant percentage (1-2%) of the F1 offspring with abnormalities. These effects were subsequently dissected out and shown to be a consequence of altered Hsp90 function. This was confirmed genetically and pharmacologically (Rutherford and Lindquist, 1998).

The genetic approach involved crossing different heterozygous *Hsp83* flies, followed by examination of the resultant offspring for phenotypic abnormalities. An increased incidence and level of severity of abnormalities was seen.

Pharmacological confirmation of the role of Hsp83 was derived from experiments in which flies were fed on food containing the Hsp90 inhibitor geldanamycin. A significant proportion (7.7%) of the flies fed with the drug spiked food had morphological abnormalities similar to those seen previously.

These authors developed an interesting hypothesis to explain the developmental abnormalities resulting from impaired Hsp90 function. This involved Hsp90 normally functioning to mask cryptic genetic variation by suppressing the expression of genetic variations affecting developmental pathways. Impaired Hsp90 function would consequently allow these variations to manifest themselves at a much higher rate in the population. It was further proposed that this could be a potential mechanism for "evolvability", as Hsp90 has a key role in many processes and developmental pathways. Perturbation of Hsp90 functions could therefore allow normally silent mutations to be expressed in the population.

5.3.4

Hsp 90 as a general molecular chaperone

The chaperone activity of Hsp90 has been investigated *in vitro* principally by the use of three assay systems. One assay system relies on examining Hsp90 protein activity alone, the others require participation of a number of Hsp90 co-chaperones. The assay involves characterising the ability of Hsp90 to prevent aggregation of one or more of a number of denatured proteins including citrate synthase, β-Galactosidase and rhodanese. *In vitro* Hsp90 alone exhibits limited chaperone activity with a wide variety of substrates. Chaperone activity is defined as the ability to maintain proteins in a folding competent state and prevent further non-specific aggregation whilst having no intrinsic refolding activity (Freeman and Morimoto, 1996; Jakob et al., 1995; Csermely et al., 1998).

Refolding of denatured test proteins including firefly luciferase has been used as a more stringent assay to determine the "active" chaperoning ability of Hsp90. Hsp90 facilitates refolding of proteins in co-operation with a number of co-chaperones including Hsp70, Hsp40, Hip, Hop, the TPR containing proteins and p23. All these proteins have some chaperoning activity *in vitro* apart from Hip and Hop, leading to a very complex system from which it is very difficult to dissect out the contribution of individual components to the resulting multi-protein chaperone machinery.

The *in vivo* chaperone activity of Hsp90 was investigated in an elegant series of experiments. Nathan and co-workers investigated Hsp90 function with steroid receptors and protein kinases (Nathan and Lindquist, 1995; Nathan et al., 1997). Temperature sensitive mutations of *S. cerevisiae* Hsp90 were used to investigate its *in vivo* functions. These resulted in a complete and rapid loss of Hsp90 activity following exposure to non-permissive temperatures. These authors suggested the role of Hsp90 *in vivo* was limited to acting predominantly to effect the maturation of a group of proteins whose folding was inherently problematic and also in facilitating the reactivation of damaged proteins. Chaperone sites in both N- and C-terminal domains of Hsp90 have been identified, differing in their substrate specificities and ATP dependence (Young et al., 1997; Scheibel et al., 1998). Functional association between the two chaperone domains was also suggested (Marcu et al., 2000).

5.3.5

Hsp90 and steroid hormone receptors

Steroid hormone receptors are soluble intracellular proteins that, dependent upon their state of activation, can shuttle between the cytosol and the nucleus, leading to activation of transcription. Hsp 90 acts as part of a dynamic multi-protein complex in the assembly and maintenance in an "activatable" state, of steroid hormone receptor complexes. The role of this family of proteins is thought to be largely to fold the hormone-binding domain of the receptor into a high affinity steroid hormone binding conformation. The interactions between Hsp90 and partner proteins in this system are the best characterised to date, although there is as yet little understanding of the mechanism involved or detail available on the interactions between individual component proteins in this multi-protein chaperone hetero-complex.

5.3.5.1

Maturation of Steroid Hormone Receptors (SHR's) is a three-step process

The process of maturation of SHR's into conformationally mature molecules can be broken down into approximately three stages of interactions with different components of the chaperone machinery, early, intermediate and mature.

In the early stages the SHR associates with Hsp70, Hsp40 and Hip, binding of Hsp90 and Hop to this pre-formed complex leads to the intermediate state. Studies have shown the interaction between Hsp70, Hop and Hsp90 to be a crucial step in this process, with Hop interacting physically with both Hsp70 and Hsp90, perhaps modulating the activities of these proteins (Chen and Smith, 1998; Johnson et al., 1998; Chen et al., 1998). Formation of the mature complex is brought about by the interaction of another accessory protein, p23, with the receptor and one of a number of TPR containing immunophilins. Hormone binding, through the high affinity hormone binding domain of the receptor, results in disassociation of the SHR/Hsp90 multi-component complex, followed by dimerisation of the receptor with subsequent transport into the nucleus. Movement into the nucleus is followed by the receptor binding to specific DNA sequences, allowing initiation and regulation of transcription. Figure 44 provides a diagrammatical representation of the component proteins and stages involved in the maturation of a generic SHR.

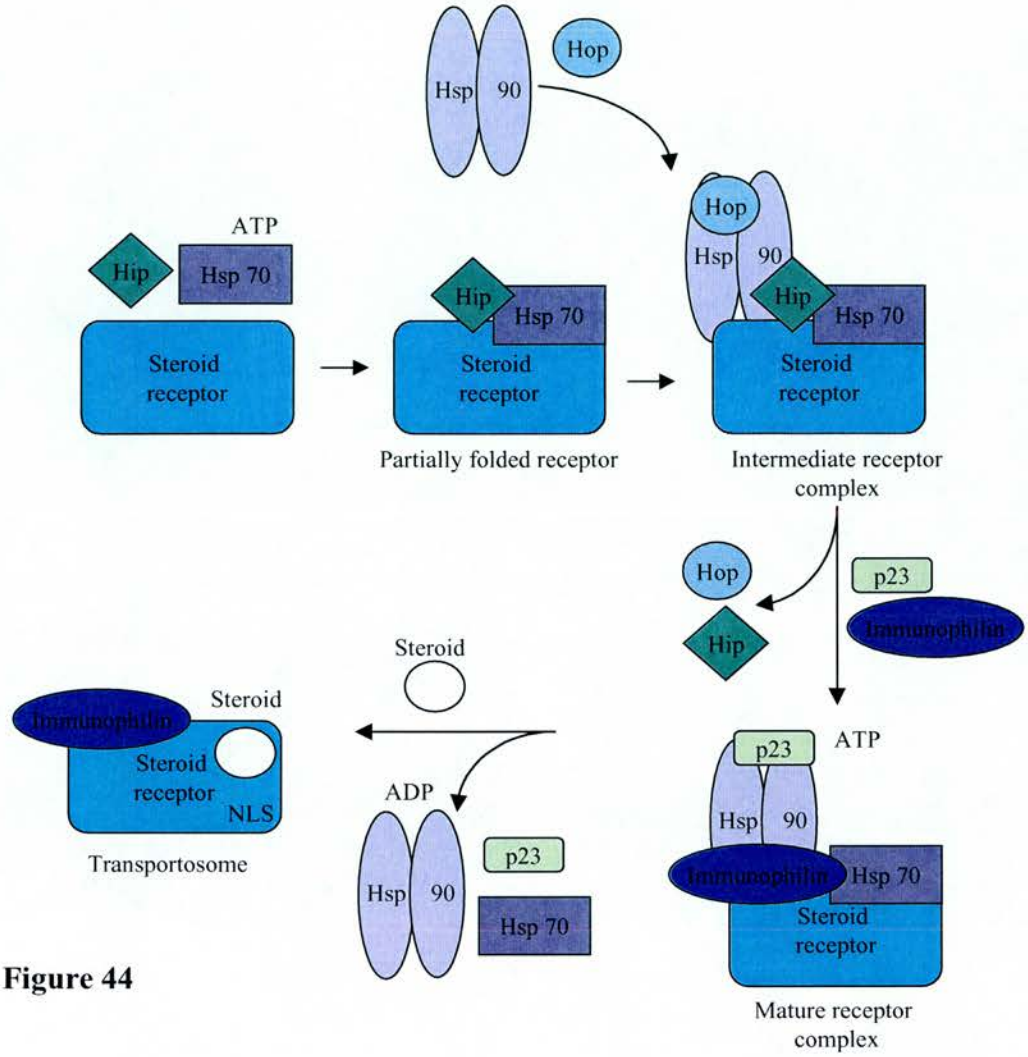


Figure 44

Illustration of the three principal stages in the maturation of a generic steroid hormone receptor. The process can be broken down into three broad stages, early, intermediate and late. Each of the stages is characterised by the presence of a complex formed from a receptor specific subset of proteins with different protein components of the chaperone machinery. The late stage is characterised by the association of the SHR with Hsp90, p23 and PP5 or one of the large immunophilins FKBP51, FKBP52 or Cyclophilin 40. The immunophilins interaction with Hsp90 is mediated through a common TPR acceptor site. Figure 44 modified from (Kimmins and MacRae, 2000).

5.4

MATERIALS AND METHODS

The C-terminal domain of Hsp90

The interaction site of Hsp90 with large immunophilins has been shown to reside in the C-terminal region of Hsp90. The experimental component of this chapter will focus on experiments utilising deletion mutant constructs of the C-terminal domain of Hsp90. The results presented will include discussions of studies carried out in an attempt to further characterise Hsp90 C-terminal domain fragments, including expression, purification and crystallisation trials

Selection of different lengths of Hsp90β deletion mutants for study

Earlier work by Carrello et al (Carrello et al., 1999), using a two hybrid system screen of a mouse cDNA library, had identified the Cyp 40 binding domain in Hsp90. All interactions between these two proteins were mediated by the carboxyl terminus of Hsp90, and were confirmed *in vitro* by bio-affinity interaction studies. The region in Hsp90 responsible for the interaction with Cyp 40 was further characterised, revealing the key residues to lie within a 124 amino acid C-terminal segment of Hsp90. A number of constructs were produced, covering this region.
The authors went on to describe the effect of deletion of the 5 extreme C-terminal residues of Hsp90 (MEEVD), which resulted in complete disruption of binding to Cyp 40.

5.4.1

Hsp 90 cloning and expression

Cloning was carried out by the Ratajczak group, Nedlands, Western Australia. The following is provided for completeness only.

To generate N-terminal His tagged expression plasmids the following cloning strategy was followed. Expression plasmids for Hsp 90β deletion mutants

incorporating codons 530-724, 581-724, 600-724 and 530-719 were constructed as follows. Each construct was PCR amplified from pET-15b-Hsp90β full length cDNA template using *Taq* DNA polymerase with primers that introduced an NdeI site at the 5'-end and a termination codon with an NdeI site at the 3'-end. The PCR fragment was ligated into pGEM-T and was then excised with NdeI digestion from a suitable clone confirmed by sequence fidelity. The gel purified fragments were cloned into the pET 28a(+) vector that had been linearised with NdeI to give the expression plasmids pET 28a(+) 530-724β, pET 28a(+) 581-724 β, pET 28a(+) 600-724 β, pET 28a(+) 530-700 β and pET 28a(+) 530-719 β.
A similar strategy was followed to produce the expression plasmid pET-28a(+) 589-732α, a construct for co-expression of Cyp 40 and Hsp90 600-724 was also prepared.

Hsp90 530-724 and 600-724 were chosen as initial targets for further characterisation. These constructs were selected to provide examples of the putative dimeric form of Hsp90, (530-724 includes the dimerisation domain of Hsp90) and monomeric protein (600-724). Preliminary expression trials also indicated these constructs expressed well, producing soluble protein, important for characterisation and interaction studies.

5.4.2

pET expression of recombinant Hsp90 in E.coli.

The standard protocol developed throughout the course of this project, for large-scale over-expression of proteins under control of the T7 promoter was used (see Chapter 2, Section 2.3) for both constructs. All work was carried out either on ice using chilled solutions, in equipment pre-cooled to 4°C, or at 4°C in the cold room.

5.4.3

Purification of recombinant Hsp90

All constructs contained an N-terminal 6-His tag to simplify purification of recombinant protein. The method to purify His tagged proteins under native

conditions, as outlined in the Qiagen Handbook was followed, and is included below for completeness.

The cell pellets were defrosted on ice and solubilised in lysis buffer containing 50mM NaH_2PO_4, 300mM NaCl, 10mM imidazole at pH 8.0, at a ratio of 2-5ml per gram wet weight of cell pellet. Lysozyme (ICN/FLOW) was added to 0.1% w/v and incubated on ice for 1 hour. An equal volume of deionised water at 4°C was added and the cell suspension mixed. The lysate was centrifuged to remove cell debris, the resulting clarified lysate was frozen at –20°C.

Clarified cell lysate was applied to a 10ml column of Ni-NTA Superflow in buffer A (50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole) at pH 8.0. The column was washed until a baseline was established, followed by elution of the tagged protein in a series of steps of increasing concentration of buffer B (50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole) at pH 8.0. Column fractions were tested on 15% SDS-PAGE gels and fractions containing the protein of interest were pooled.

5.5
CHARACTERISATION OF HSP90

5.5.1
Dynamic light scattering analysis of the purified protein

Dynamic light scattering studies of the purified protein were carried out as described in Chapter 2, Section 2.4.1.

5.5.2
Electrospray ionisation mass spectrometry

The proteins were further characterised by electrospray ionisation mass spectrometry on a SELDI (Surface enhanced laser desorption/ionisation) spectrometer (Ciphergen). This instrument analyses samples applied directly to a ProteinChip. Sample application to an appropriate ProteinChip can bring about enrichment of the desired protein, desalting and also provides a sample for mass spectrometry analysis. The SELDI instrument has an interface to a quadropole time-of-flight tandem mass spectrometer. Thanks to Dr. Andrew Cronshaw for the SELDI results.

5.5.3
Crystallisation trials of Hsp90

The hanging drop vapour diffusion method as outlined in Chapter 2, Section 2.5.1.3 was used throughout this series of experiments.
Crystallisation trials were carried out on both Hsp90 530-724 and 600-724 protein fragments.

5.6

RESULTS AND DISCUSSION

5.6.1

Purification of Hsp90 530-724

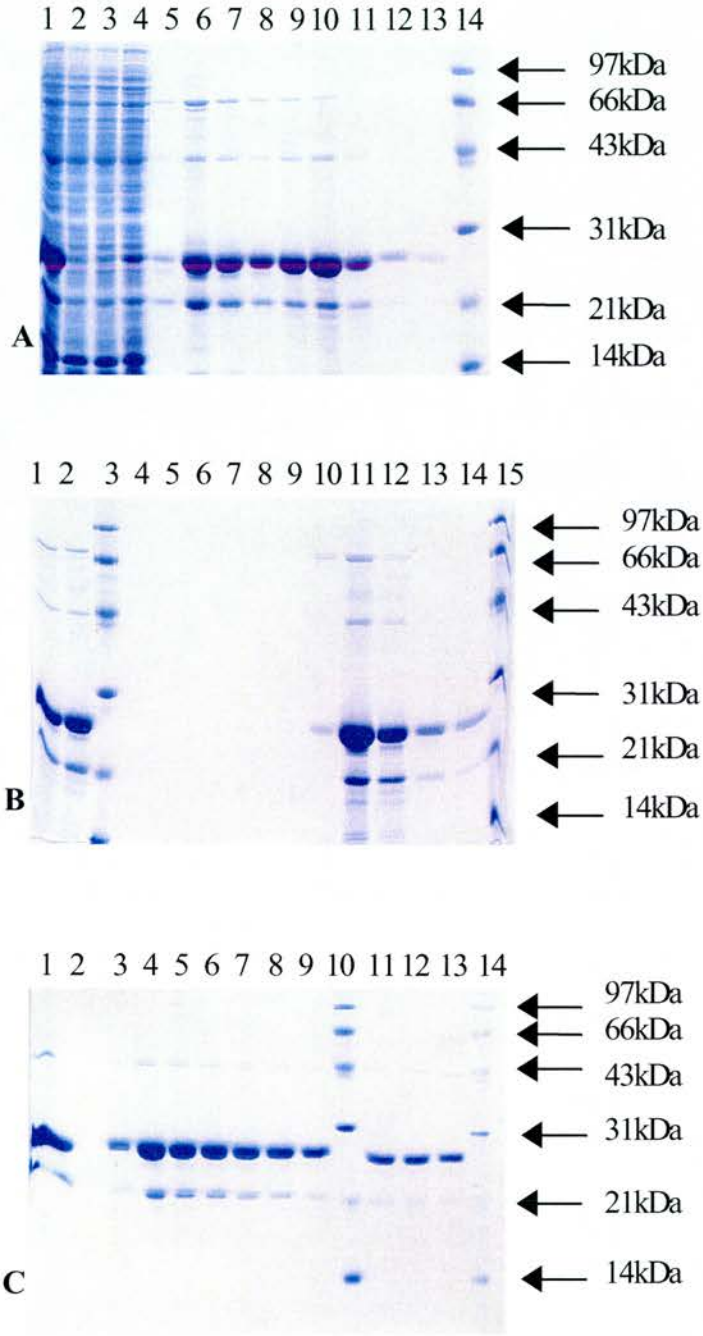


Figure 45

A representative preliminary purification run of Hsp90 530-724.

(A)
A representative gel illustrating over-expression and preliminary purification of Hsp90 530-724 on Ni-NTA superflow. Lane 1, clarified cell lysate. Lanes 2 - 4 wash fractions, lanes 5-7 elution with 50mM imidazole. Lanes 7 - 13 elution with a gradient of imidazole to 200mM. Lane 14, molecular weight markers

(B)
Pool of fractions from lanes 5-13, applied to Mono Q column. Lanes 1 and 2, "start material" post nickel column, unfiltered and filtered respectively. Lanes 4-7, wash fractions, lanes 8-14 elution of proteins by an NaCl gradient to 500mM. Lanes 3 and 15, molecular weight markers.

(C)
Pool of fractions post elution from nickel column, concentrated and applied to a gel filtration column (Sephacryl S-200). Lane 1, "start material", lanes 2-9 and 11-13 sequential fractions from peaks on the chromatogram. Lanes 10 and 14, molecular weight markers.

5.6.2

Purification of Hsp90 600-724

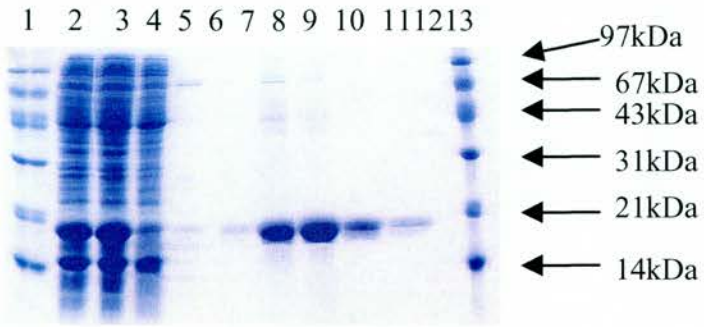


Figure 46
A representative preliminary purification run of Hsp90 600-724

A representative gel illustrating over-expression and initial purification trials of Hsp90 600-724 on Ni-NTA Superflow. Lane 1, molecular weight markers, lanes 2 and 3 clarified cell lysate unfiltered and filtered respectively. Lanes 4 and 5, wash fractions, lanes 6-12 elution fractions with 50mM imidazole. Lane 13 molecular weight markers.

Protein standards providing molecular weight calibration were run on each gel, allowing direct comparison between the proteins. SDS PAGE analysis revealed the Hsp90 protein fragments running with a higher than expected molecular weight. The estimate of molecular weight provided by analysis of the electrophoretic mobility of proteins can be influenced by a number of features, including the protein construct (discussed in greater detail later), the charge and nature of the proteins under test and the extent to which these proteins bind the solubilisation agent SDS. Very acidic proteins have been found to bind small amounts of SDS, the predicted pI for both constructs investigated to date is less than 5, perhaps contributing to their greater than expected estimated molecular weight.

Hsp90 530-724

As discussed above, the predominant protein produced by this construct is one of approximately 26kDa. The over-expression system produces a relatively high yield of protein that could be enriched to greater than 90% pure after one nickel agarose resin affinity chromatography step. Subsequently, increased levels of purification have proved harder to achieve as can be seen from the SDS PAGE gels pictured above (Figure 45). Application of the protein to high resolution anion exchange

chromatography (Mono Q resin) produced little or no improvement in protein purity, as may have been hoped had the contaminants differed greatly in their charge from Hsp90 530-724. Gel filtration also failed to separate the target protein from the principal contaminants, although the separation required was approaching the limit of resolution for the Sephacryl S-200 column. All proteins eluted very early, suggestive of target protein and contaminants being present as a high molecular weight aggregate. These results suggest the sample was not in an optimum state for crystallisation trials.

Hsp90 600-724

Figure 46 above illustrates the one step purification protocol employed for Hsp90 600-724. SDS PAGE suggested the protein was greater than 95% pure after a single nickel agarose affinity chromatography step. No further steps were taken to further purify this construct, although trace higher molecular weight contaminants could be seen on the gel.

5.6.3

Dynamic light scattering analysis

Table 14 illustrates a typical dataset obtained from a dynamic light scattering experiment carried out using both Hsp90 constructs purified as described above. As outlined in Chapter 2, measurements were taken at room temperature using a DynaPro 801 instrument with micro-sampler attachment.

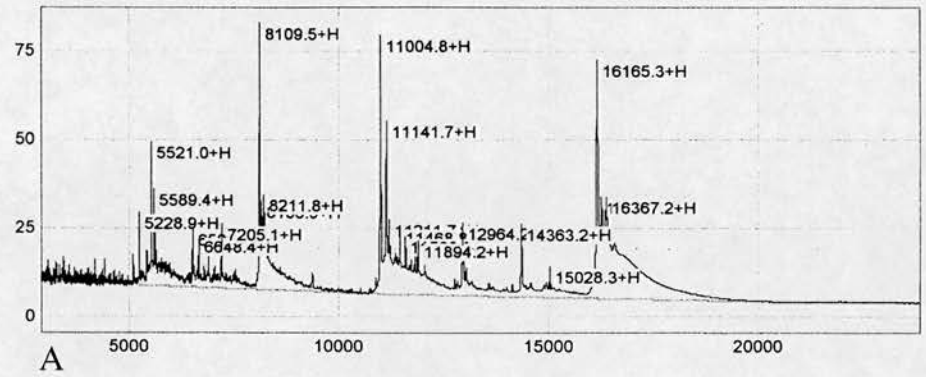
Measurement number	**Polydispersity (nm)**		**Estimated molecular weight in kDa**		**Baseline**		**SOS Error**	
Protein fragment length	**600-724**	**530-724**	**600-724**	**530-724**	**600-724**	**530-724**	**600-724**	**530-724**
1	1.7	1.5	214	110	1.002	1.006	2.263	3.734
2	1.3	1.0	216	100	1.000	1.004	2.376	3.596
3	1.7	3.0	227	151	1.002	1.012	2.181	9.884
4	2.3	1.6	254	99	1.007	1.003	4.721	3.231
5	2.8	2.7	279	151	1.035	1.041	4.974	7.662
6	1.8	1.7	317	107	1.016	0.998	6.615	4.436
7	1.9	1.6	276	98	1.008	1.003	4.219	2.736
8	1.5	1.2	234	107	1.002	1.001	2.735	4.094
9	2.2	1.0	247	105	1.002	1.002	3.094	2.899
10	2.1	0.7	246	114	1.002	0.999	2.321	3.122
11	2.1	2.3	225	179	1.002	1.028	2.128	10.69
12	1.2	1.8	244	108	1.003	1.005	2.934	4.639
Mean values	**1.9**	**1.7**	**248**	**119**	**1.007**	**1.008**	**3.382**	**5.060**

Table 14

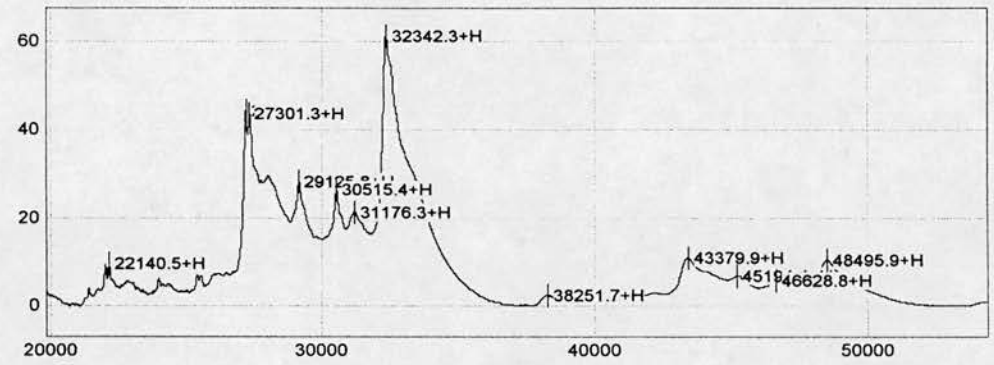
These results suggested the proteins were not monodisperse but were in fact present as multimers/aggregates. The results for Hsp90 530-724 are in broad agreement with those of the gel filtration, suggesting the protein is present in solution as a high molecular weight aggregate.

5.6.4

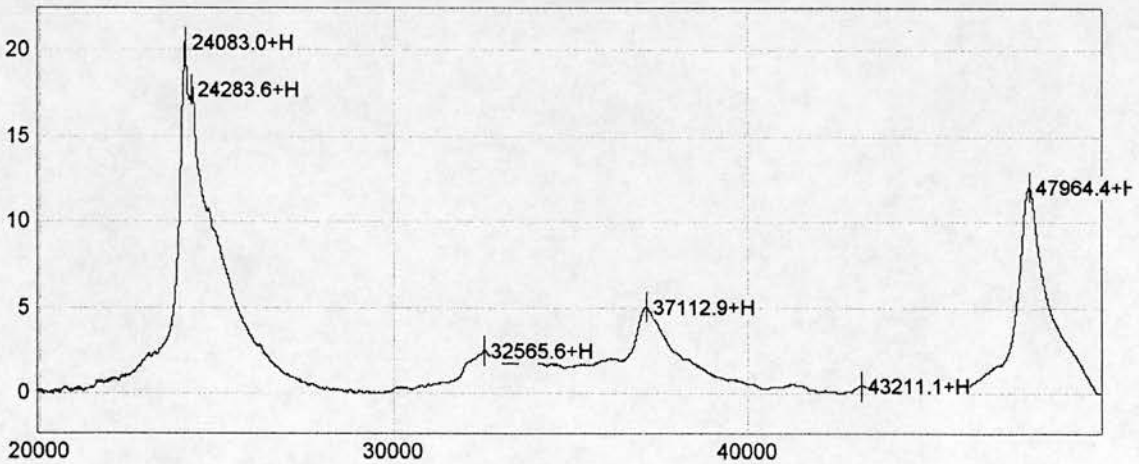
Mass spectrometric analysis of Hsp90 530-724 and 600-724



A



B



C

Figure 47

Results from electrospray ionisation mass spectrometry on a SELDI (Surface enhanced laser desorption/ionisation) spectrometer of two Hsp90 C-terminal fragments. Panels A and B results for the Hsp90 600-724 sample, panel C Hsp90 530-724. The samples can be seen to contain a large amount of contaminants.

Protein	Theoretical MWT (Da)	MWT of tag (Da)	Combined total	Major peak(s) in mass spectrometry results
Hsp90 600-724	13,978Da	2181Da	**16,159Da**	**16,165Da,** 8,110Da, 11,005Da, 16,165Da 27,301Da, 32,342Da
Hsp90 530-724	21,897Da	2181Da	24,078Da	**24,083Da, 24,283Da 47,964Da**

Table 15

Summary of results from mass spectrometry experiments. The theoretical molecular weight of the protein sequences were calculated by the Compute pI/Molecular Weight tool at the ExPASy server (Wilkins M R and Gasteiger E, 1998), mass spectrometry results were obtained as outlined earlier in this chapter in Section 5.5.2. The results suggest strongly that Hsp90 530-724 would appear to be a better candidate for structural studies than the shorter fragment, which appears to be present in numerous isoforms.

The hexa-His tagged vector used in the cloning of these fragments of Hsp90 was included to facilitate purification of the recombinant protein. The cloning/expression region of the pET-28a-c(+) is shown below in Figure 48. This vector does, however, add an estimated 2181 daltons to the final molecular weight of the protein products.

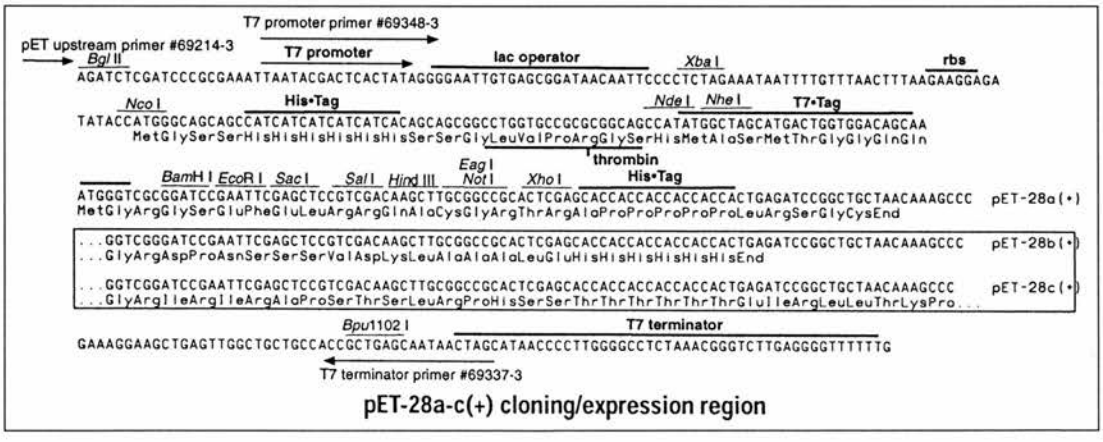


Figure 48

Cloning and expression region of the pET-28a-c(+) vector. Figure taken from Novagen technical reference literature.

Taking into account the additional residues from the vector, results from the mass spectrometry confirm that proteins of the expected molecular weights are being produced. The relative proportions of desired product to contaminants appears to be disappointing in the case of the shorter construct.

The results for the longer fragment (530-724) were more readily interpretable than those of the shorter fragment, both sets of results will now be considered in greater detail.

Hsp90 600-724

Hsp90 600-724 is apparently very heterogeneous, numerous peaks being identified from the sample. This included a small proportion at approximately the expected molecular weight of the intact protein (16,159Da). This was a somewhat unexpected result as the sample appeared to be relatively highly purified, when analysed by SDS PAGE and the proteins visualised by coomassie blue staining (Figure 46). Higher molecular weight contaminants could be seen in the sample however no significant amount of smaller contaminants were detectable on the gels. The presence of a significant proportion of lower molecular weight contaminating proteins in the sample suggested the protein was prone to proteolytic breakdown.

Secondary structure predictions for the C-terminal domain of Hsp90 (to be discussed in greater detail in Chapter 6) suggest the molecule is largely helical, with short regions of strand and the final 38 C-terminal residues lacking predicted secondary structure. A helix running from residue 598-608 is predicted. Creation of a deletion mutant from residue 600-724 could therefore be disrupting one of the helices of the protein fragment. This potential disruption of secondary structure elements, together with the unstructured C-terminal tail of the protein suggests great potential for inducing structural instability and susceptibility to proteolytic breakdown into the protein, as a possible consequence of incorrect folding.

Mass spectrometry analysis was carried out some time after the protein was purified, there was therefore opportunity for the protein to have degraded over time. Alternatively, the contaminants seen may have been as a consequence of the action

of proteolytic enzymes, the protein having been prepared and stored without the addition of proteolytic inhibitors.

Hsp90 530-724

This particular length of construct was chosen as it included the dimerisation domain of Hsp90, located within the C-terminal region of the protein (residues 524-724), as identified from studies using various lengths of deletion mutants (Nemoto et al., 1995, Carrello et al., 1998).

The results for Hsp90 530-724 suggest two of the three major peaks present in the sample correspond reasonably well to the expected molecular weights of both monomeric and dimeric forms of the protein. Structural studies of both monomeric and dimeric Hsp90 fragments are planned. The apparent presence of two monomeric isoforms of the protein, separated by approximately 200 daltons only, is however, potentially problematic. Application of the purified protein to a suitable gel filtration column should theoretically be capable of resolving the protein into its monomeric and dimeric forms, however it would have no effect upon the two monomeric isoforms. As discussed earlier in this chapter in Section 5.6.2, gel filtration failed to resolve monomeric Hsp90 530-724 from contaminating proteins, should the interaction prove to be sufficiently robust gel filtration would be the method of choice to fractionate the protein into monomeric and dimeric forms.

5.6.5

Crystallisation trials of fragments of the C-terminal domain of Hsp90

The proteins were concentrated to 20mg/ml, the following crystallisation trials were set up for both constructs

- Crystal screen I
- Crystal screen II
- Crystal Structure Screen I
- Crystal Structure Screen II
- Ammonium sulphate
- Polyethyleneglycol 4000

To date no protein crystals have been obtained from any of these conditions. A number of conditions looked more promising, these included ammonium sulphate which produced "promising precipitate". Further refinement of conditions including a broad pH screen (pH 2.2 – 9.0), suggested lower pH looked more promising.

5.7
DISCUSSION

The results presented in this Chapter include those from preliminary large-scale expression, purification and characterisation of two constructs of the C-terminal domain of Hsp90. The shorter construct, 600-724 was chosen to produce monomeric Hsp90, 530-724 contains the dimerisation domain of the protein. Analysis of the different constructs therefore would allow further dissection of the C-terminal domain of Hsp90. Availability of four C-terminal constructs of Hsp90β covering residues 530-724 of the protein domain implicated in interactions with the immunophilins including Cyclophilin 40, provides an opportunity to gain some further insight into both inter and intra molecular protein/protein interactions of this key regulatory molecule.

The approach taken in these preliminary experiments was to attempt purification and crystallisation of the C-terminal domain of the Hsp90 molecule alone. The results presented earlier for both constructs suggest the protocol for the extraction, purification and handling of these proteins required some optimisation to increase the chances of obtaining useful crystals.

The C-terminal His tag used in this series of constructs added approximately 2181Da to the proteins produced, and may be influencing the efficient folding of the proteins. The tag used in these constructs is cleavable. Removal of these tags by thrombin cleavage would provide a different protein species to be used in structural studies. However, as discussed in Chapter 2, proteolytic cleavage of target proteins has its own hazards, including generation of heterogeneity and inappropriate/inefficient cleavage of target proteins. Should the tagged proteins prove resistant to crystallisation, removal of the tag could be attempted.

The longer construct appears to be vulnerable to aggregation, perhaps through the formation of non-specific disulphide bonds between available cysteine residues (there are 3 within the construct). To address this question, the protein could be purified and stored under more strongly reducing conditions than those used previously.

The presence of micro-heterogeneity within a protein sample can be problematic for crystallisation. This may prove to be the case with Hsp90 530-724. Further experiments will be necessary to determine if the protein peaks present at approximately 24,083 and 24,283Da are genuinely present in roughly equivalent amounts, and also whether they can be resolved prior to crystallisation experiments.

Mass spectrometry results suggest a significant proportion of the sample contains an Hsp90 dimer, additional experiments will be required to optimise and enhance the formation of dimeric Hsp90, producing a dimer robust enough to be maintained throughout purification and crystallisation processes.

The shorter construct used in these experiments appears to be prone to both aggregation and degradation. Steps to minimise deleterious effects upon the protein might involve addition of a proteolytic inhibitor cocktail to the buffers used during purification, hopefully minimising protein degradation. Proceeding as quickly as possible from cell lysis to crystallisation trials may also be crucial, anecdotal evidence suggests "fresh" protein can be very important in helping tip the balance in favour of crystallisation for troublesome proteins. The apparent aggregation seen from both mass spectrometry and light scattering experiments is also problematic. This construct does not contain any cysteine residues, however as discussed earlier in the chapter a predicted helix is disrupted by this construct, perhaps leading to instability of this particular construct. One approach that might yield more stable protein could involve re-engineering the construct, taking into account predictions of secondary structure.

CHAPTER 6
MODELLING STUDIES

CHAPTER 6A
COMPARATIVE PROTEIN MODELLING STUDIES OF ELEVEN CYCLOPHILIN 40 HOMOLOGUES

6.1
INTRODUCTION TO COMPARATIVE PROTEIN MODELLING

Comparative protein modelling is based upon the observation that proteins with highly similar sequences have similar structures (Chothia and Lesk, 1986). In the process of comparative protein modelling, a known 3D structure of a related family member is used as the basis for extrapolating a new structure from a protein sequence, reviewed in (Bajorath et al., 1993).

A search of the SWISS-PROT protein sequence database identified eleven proteins with a high degree of homology to bovine Cyclophilin 40. Bovine Cyp 40 remains the only family member to be structurally characterised. The aim of this part of the project was to derive modelled structures of each of the other family members, utilising the high resolution X-ray structure of Cyp 40 as a template. In an effort to understand the function of a protein, knowledge of its shape and where and how it interacts with protein substrates would be very useful information.

Analysis of modelled structures of these proteins might have provided insight into the roles played by these proteins in their particular systems, possibly shedding light onto the fine tuning of response made possible by minor changes in amino acid sequence and hence structure. The aim of this part of the project was to compare and contrast the experimentally derived crystal structure with modelled structures of the other family members.

6.2
MATERIALS AND METHODS

6.2.1
The SWISS-MODEL server for automated comparative protein modelling

The automated internet based SWISS-MODEL server (Guex and Peitsch, 1997; Guex et al., 1999) was used in an attempt to produce modelled structures of the eleven cyclophilin 40 homologues isolated from the SWISS-PROT database. The SWISS-MODEL programme has an absolute requirement for at least one sequence of known 3-D structure that has significant sequence similarity (ideally greater than 40%) to the target protein.

The SWISS-MODEL programme compares the target sequence with a database of sequences derived from the Brookhaven PDB, using both FASTA and BLAST sequence comparison programmes. Selection of suitable template structures is additionally restricted to those proteins that share at least 35% residue identity with 40% of the target sequence, as determined by an algorithm within the programme. This process of screening proteins for the most highly homologous allows selection of several of the potentially most suitable templates for a given target sequence. Up to 10 templates can be used during the modelling process. A "reference" template is selected from those available, and is usually the protein with the highest sequence similarity to the target. All other selected templates are superimposed upon the reference template. 3-D matching of the structures results from superimposing the corresponding Cα atom pairs selected automatically from the highest scoring local sequence alignment, generating a structurally correct (multiple) sequence alignment.

6.2.2

"First approach mode"

First approach mode provides a very simple interface to the SWISS-MODEL programme. The programme was accessed via an internet site http://www.expasy.ch/SWISS-MODEL.html. The Cyp 40 protein sequences (Table 16) either as a SWISS-PROT AC code number or in FASTA format, were individually supplied to the programme with optional template files. All results were returned *via* e-mail. The modelling process took approximately 15 minutes to return most results, however on occasions the modelling request was unsuccessful. A second mode was available with which to retry automated modelling, the so-called "Optimise mode".

6.2.3

"Optimise mode"

Optimise mode requires use of another related programme, the Swiss-PdbViewer, (http://www.expasy.ch/spdbv/), (Guex and Peitsch, 1997). The Swiss-PdbViewer provides an interface that allows superposition of two or more protein sequences, allowing generation of structural alignments between template and target molecules. Additionally these molecules can be viewed and manipulated with the goal of improving alignment between template and target molecule, prior to resubmission of the modelling request to SWISS-MODEL.

6.2.3.1

Submission and modelling of target sequences

The amino acid sequence of the protein to model was loaded into the programme. The target molecule is initially represented in the graphical interface window as a single long helix. The PDB file of structure of bovine Cyp 40 was used throughout this series of modelling requests as the template molecule, and was also loaded into the graphical interface. The target sequence was then mapped onto the template,

resulting in an automatically generated structural alignment between the two molecules. This alignment can be manipulated either automatically or manually (in a separate alignment window), the modelled structure being automatically updated as changes are made in the alignment between the two proteins. For all target proteins the sequence alignments were optimised by altering the automatic alignment generated by Swiss-PdbViewer to conform to that generated by individual ClustalW (Higgins et al., 1996) alignments between each of the target molecules and bovine Cyp 40. Following generation of an optimised alignment the modelling request was re-submitted to SWISS-MODEL.

6.2.3.2

Visualisation and analysis of modelled structures

Results from the modelling process were returned *via* e-mail and included an alignment of the amino acid sequences of the two proteins, together with the co-ordinates for the modelled structure as a PDB file. The structures were visualised using WITNOTP (Widmer, 1997). The two structures were firstly separated into molecule A, (Cyp 40, supplied template) and molecule B, (target molecule), each being saved as a separate PDB file, allowing easier manipulation of individual proteins within this program.

The following approach was routinely taken to analyse structures generated by SWISS-MODEL. Following separation of the two structures (template and target), each protein was coloured, magenta being reserved in each case for the template structure (bovine). The molecules were initially displayed without side-chains, to facilitate identification of areas of divergence from the template. This "whole molecule" image was captured and saved, and representative residues of particular interest were labelled. This was followed by more in depth examination of both molecules, particularly around areas highlighted by the initial analysis. Side-chains were added later to both molecules.

Protein Ac no.	Organism	Length (aa)	% Identity	Reference
Cyp 40 P26882	B. taurus	370	100%	(Taylor et al., 2001)
Cyp 40 Q08752	*H. sapiens*	370	94%	(Kieffer et al., 1992)
Q9CR16	*M. musculus*	370	92%	(Kawai et al., 2001)
Cyp 40 Q9C566	*A thaliana*	361	48%	(Berardini et al., 2001)
Wis2^{+} Q11004	*S. pombe*	356	43%	(Weisman et al., 1996)
Cpr6 P53691	*S. cerevisiae*	371	46%	(Duina et al., 1996a)
Cpr7 P47103	*S. cerevisiae*	393	33%	(Duina et al., 1996a)
Cyp 41 Q9P3X9	*N. crassa*	375	53%	Sequence database
Q9N639	*L. major*	354	44%	Sequence database
Q9XIL2	*A. thaliana*	404	49%	Sequence database
Q9LY75	*A. thaliana*	570	42%	Sequence database
Q9VT21	*D. melanogaster*	383	39%	(Adams et al., 2000)

Table 16

A search of the SWISS-PROT protein sequence database identified eleven proteins with a high degree of homology to bovine Cyclophilin 40. These proteins are identified in Table 16 by SWISS-PROT primary accession number and by name where one had been used in the literature. Sequence identity for each of the proteins is compared with that of bovine Cyp 40 (over full length Cyp 40). Various functions have been suggested for bovine and human Cyp 40 (discussed in Chapter 4), however the majority of these proteins are largely functionally uncharacterised. Some data is available on Cpr6, Cpr7, Wis2 and the shortest Cyp 40 from *A. thaliana*, this is discussed in the following section.

6.3

THE CYCLOPHILIN 40 SUBFAMILY

6.3.1

The product of the Wis2$^+$ gene of Schizosaccharomyces pombe encodes for a member of the Cyclophilin 40 family.

Wis2$^+$ was first isolated as a multi-copy suppressor of the cell cycle defect of the triple knockout strain *wee-1 50 cdc25-22 win1-1* (Warbrick and Fantes 1992, Weisman et al., 1996). When grown on minimal media at the permissive temperature of 35°C, the *wee-1 50 cdc25-22 win1-1* triple mutant has a conditional lethal phenotype. The triple mutant cells are highly elongated, a phenotypic characteristic of cell cycle mutants, and also have abnormally bent and branched cells. Introduction of a multicopy plasmid containing the open reading frame (ORF) responsible for *Wis2*$^+$ activity into the triple mutant strain suppressed its conditional lethal phenotype.

Translation of the *Wis2*$^+$ ORF predicted a polypeptide of 356 amino acids and 40.1kDa. The translated protein was identified as a member of the Cyp 40 family. Homology between *Wis2*$^+$ and bovine Cyp 40 over the 360 amino acid overlap is 43% identical, over the TPR domain the degree of identity is lower at 31%, with the third TPR unit being most divergent. To date, *Wis2*$^+$ remains the only type 40 cyclophilin to be identified by the *S. pombe* genome sequencing project.

The 18kDa cyclophilin (*cyp 1*) from *S. pombe* was fused to the C-terminal domain from *Wis2*$^+$. This hybrid protein suppressed the triple mutant phenotype with comparable efficiency to that of "native" *Wis2*$^+$, whilst the cyclophilin domain alone had no such effect, showing the C-terminal domain to be critical in this interaction.

The effect of heat shock on *Wis2*$^+$ transcript was also investigated. A 10 to 15 fold increase in the level of *Wis2*$^+$ mRNA was observed following a shift of exponentially growing wild-type cells from 25 to 37°C. A further 2 fold increase was observed when cells were shifted from 25 to 42°C. These results were highly suggestive of a role for Wis2 in response to heat shock, however preliminary experiments showed no

differences in survival following heat shock between wild-type and Wis2Δ cells (Weisman et al., 1996).

Subsequent studies by the same group (T Gaskell, personal communication) proved the C-terminal TPR domain alone to isolate Hsp 90 from extracts of *S. pombe*. These results suggest Wis2 is exerting its effect on mitosis by modulation of the activity of an Hsp90 homologue.

6.3.2

Cpr6 and Cpr7 from Saccharomyces cerevisiae encode for members of the Cyclophilin 40 family

Two Cyp 40 type cyclophilins were isolated from *S. cerevisiae*, (Duina et al., 1996a). DNA sequence analysis showed the isolated clones to contain an ORF of 918 nucleotides (2 clones) and a different in frame ORF of 969 nucleotides. These ORF's represented portions of two new members of the cyclophilin family, Cpr6 and Cpr7. The entire Cpr6 and Cpr7 genes were cloned and characterised (Duina et al., 1996a) and encode for proteins with predicted molecular weights of 42 and 45kDa, respectively, sharing 38% sequence identity over their entire length. Of the two proteins, it was reported Cpr6 most closely resembled human Cyp 40 (44% identity over the whole protein, 58% identity over the cyclophilin domain and 31% identity over the TPR domain).

The effects of deletion of both genes, individually and in combination, were studied (Duina et al., 1996a). Cpr7Δ showed a slow growth phenotype, (the doubling time for wild type cells being 1.5 hours, relative to 2.6 hours for Cpr7Δ cells). Cpr6Δ cells had growth rates indistinguishable from those of wild type cells, whilst the Cpr6Δ Cpr7Δ double mutant exhibited the slow growth defect attributable to that seen with Cpr7Δ alone. These results showed that Cpr7 was required for normal growth in *S. cerevisiae,* however both Cpr6 and Cpr7 were non-essential.

Cpr6 and Cpr7 have also been shown to form a stable association with Hsp 90, having been identified in multiprotein complexes with Hsp90 and also from

bioaffinity interaction studies. The C-terminal domain of each of these proteins is sufficient for the interaction between the two proteins (Duina et al., 1998a)

The *S. cerevisiae* immunophilins were further studied. All 12 were knocked out individually by gene disruption, in all cases the null mutant strains were viable. In two cases, *cpr7* and *fpr1,* mutations conferred a slow growth phenotype when compared with the wild-type strain. *fpr1* is the *S. cerevisiae* homologue of FKBP12, yeast strains lacking FKPB12 are resistant to both FK-506 and rapamycin and have a slow growth phenotype (Wiederrecht et al., 1991).
To address the question of functional redundancy between members of the family different combinatorial mutants were generated (Dolinski et al., 1997), for example all five of the secretory pathway immunophilins were disrupted. None of these combinations exhibited any novel phenotype or resulted in any worsening of the single mutant phenotypes seen with *cpr7* and *fpr1*. A dodecuplet mutant strain lacking all 12 immunophilins was found to be viable (Dolinski et al., 1997). These results suggests that immunophilins did not play a general role in protein folding but probably performed highly specific functions through interactions with unique sets of restricted partner proteins as yet unidentified, possibly only under conditions of cellular stress. A number of studies had shown that transcription of several immunophilins was induced by heat shock or the accumulation of misfolded protein (Sykes et al., 1993; Partaledis and Berlin, 1993).

6.3.3

Cyp 40 type proteins from Arabidopsis thaliana, Drosophila melanogaster, Leishmania major and Neurospora crassa

Little is known of the detailed functions of the remaining bovine Cyp 40 homologues. *Arabidopsis thaliana* Cyp 40 has been implicated in the regulation of specific signalling pathways, loss of function mutants in the gene encoding for the Cyp 40 protein, results in plants with altered morphology (Berardini et al., 2001).

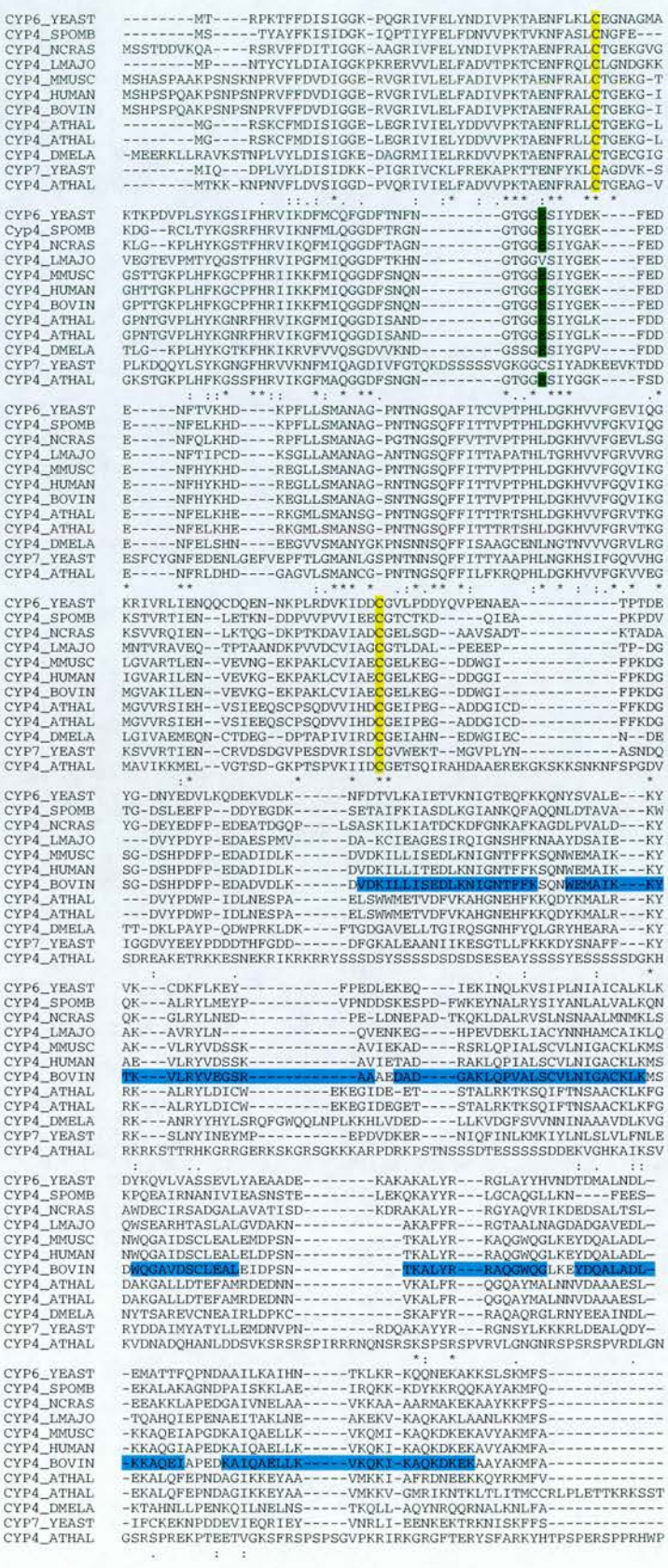


Figure 49

Alignment of the twelve Cyp 40 type proteins identified from the SWISS-PROT protein sequence database. Sequences were aligned using ClustalW (Thompson et al., 1994) with the following nomenclature CYP6_YEAST, CYP7_YEAST, respectively Cpr6 and 7 from *S. cerevisiae*. All other members of the family are represented by CYP_4 followed by an abbreviated version of the organism name. The three proteins from *A. thaliana* are numbered 1 to 3. Sequences of *A. thaliana* proteins 2 and 3 are incomplete.

A number of features (as noted in Chapter 4, section 4.6.5) are highlighted in the alignment and include conserved cysteine residues (yellow) and a highly conserved glutamate (green). The TPR helices of Cyp 40 are also delineated within the alignment (cyan).

The sequence alignment was used to help assess the validity of modelled structures generated by SWISS-MODEL.

6.4

RESULTS AND DISCUSSION

6.4.1

Structures generated by automated model building

Organism	First approach mode	Optimise mode
H. sapiens	Model generated	N/A
M. musculus	Model generated	N/A
A thaliana (361aa)	Model generated	N/A
N. crassa	Model generated	N/A
D. melanogaster	Model generated	N/A
L. major	Model generated	N/A
S. pombe *(Wis2+)*	Failed	Model generated
S. cerevisiae *(Cpr6)*	Failed	Failed
S. cerevisiae *(Cpr7)*	Failed	Failed
A. thaliana (404aa)	Failed	Failed
A. thaliana (570aa)	Failed	Failed

Table 17

Results from the automated SWISS-MODEL protein-modelling server. Eleven Cyp 40 homologues were submitted to the server. The programme generated models for six of the protein targets in "First approach" mode, one round of "Optimise" mode resulted in a modelled structure for *S. pombe* Wis2. All other targets were assumed to be insufficiently similar to the template molecule with the consequent failure of model building, despite optimisation of sequence alignments.

6.4.2

Human and mouse Cyp 40 structures

Alignments of the three proteins shows no insertions and/or deletions, indeed 346 and 339 residues out of 370 are identical for the human and mouse sequences respectively. The modelled structures of human and mouse Cyp 40 can therefore be described as essentially identical to that of bovine Cyp 40 and are not therefore not shown, having been fully covered in Chapter 4.

6.4.3

Modelled structures of five Cyp 40 homologues

Each modelled structure is presented individually overlaid with the experimentally derived X-ray structure of bovine cyclophilin 40, discussed in Chapter 4. Figure 55 in this series is an overlay of the five modelled cyclophilins with Cyp 40. Bovine Cyp 40 is coloured magenta throughout this series of figures. All figures were prepared with INSIGHT, using PDB files generated by WITNOTP.

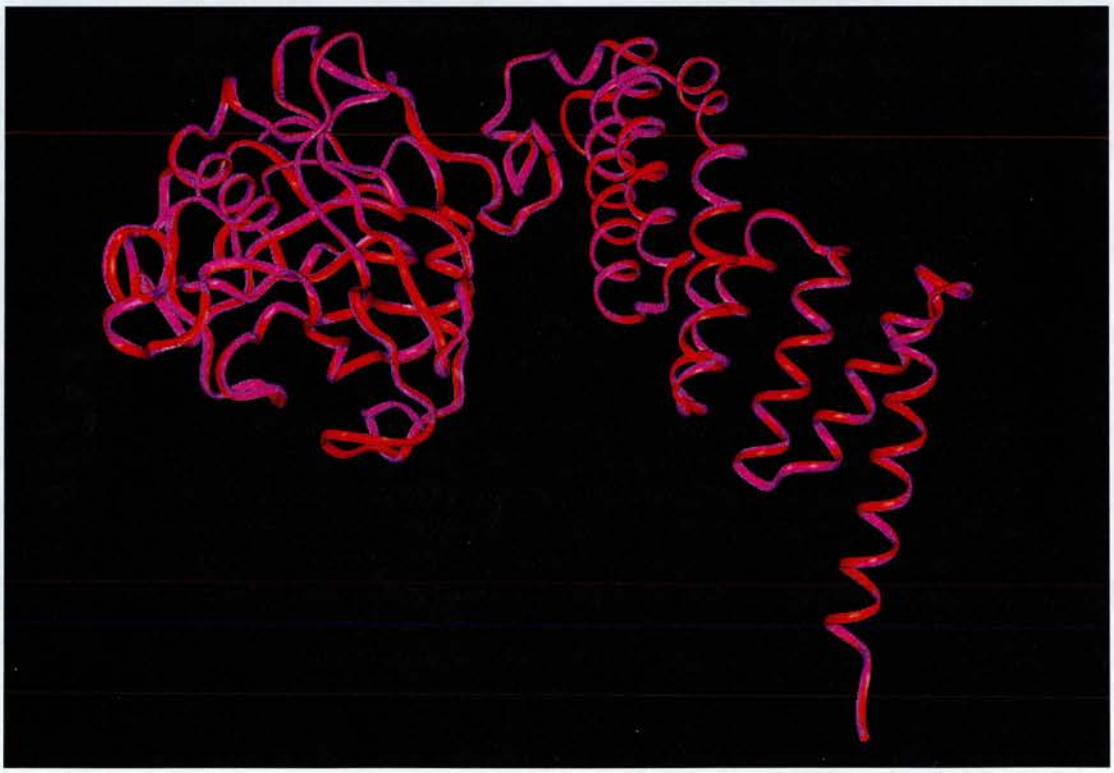


Figure 50

Overlay of the structures of *A. thaliana* Cyp 40 (red) and bovine Cyp 40 (magenta).

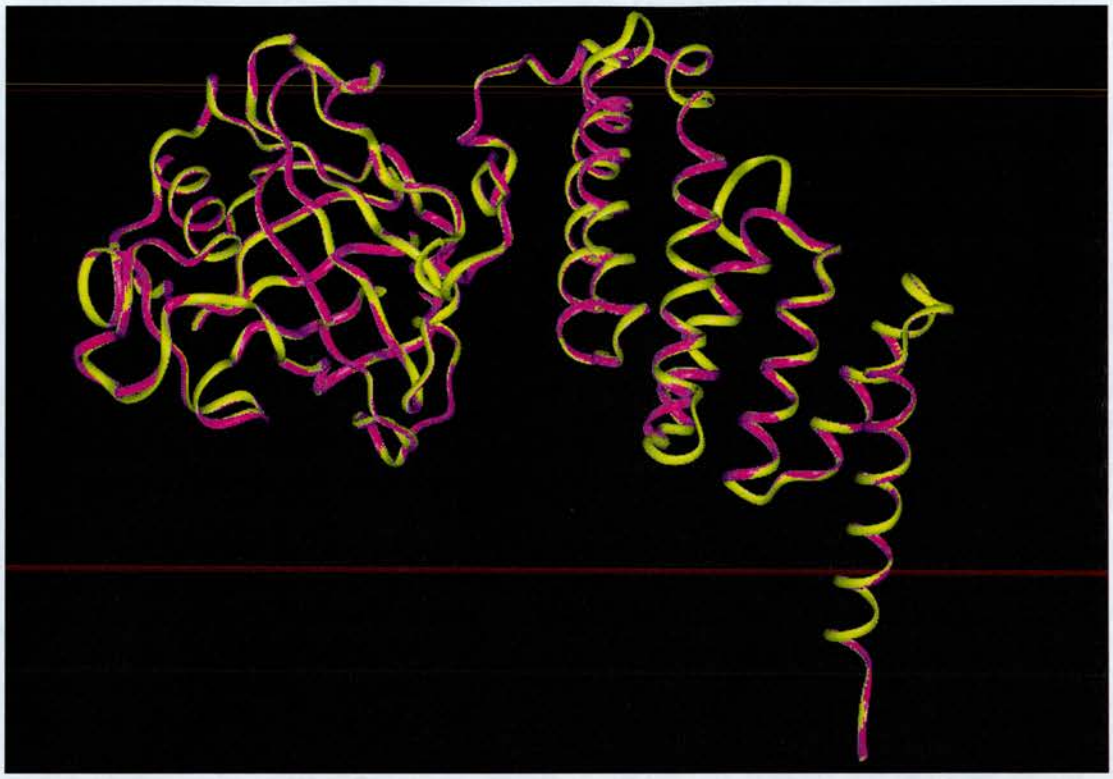


Figure 51
Modelled structure of *N. crassa* Cyp 41 (yellow) overlaid on bovine Cyp 40

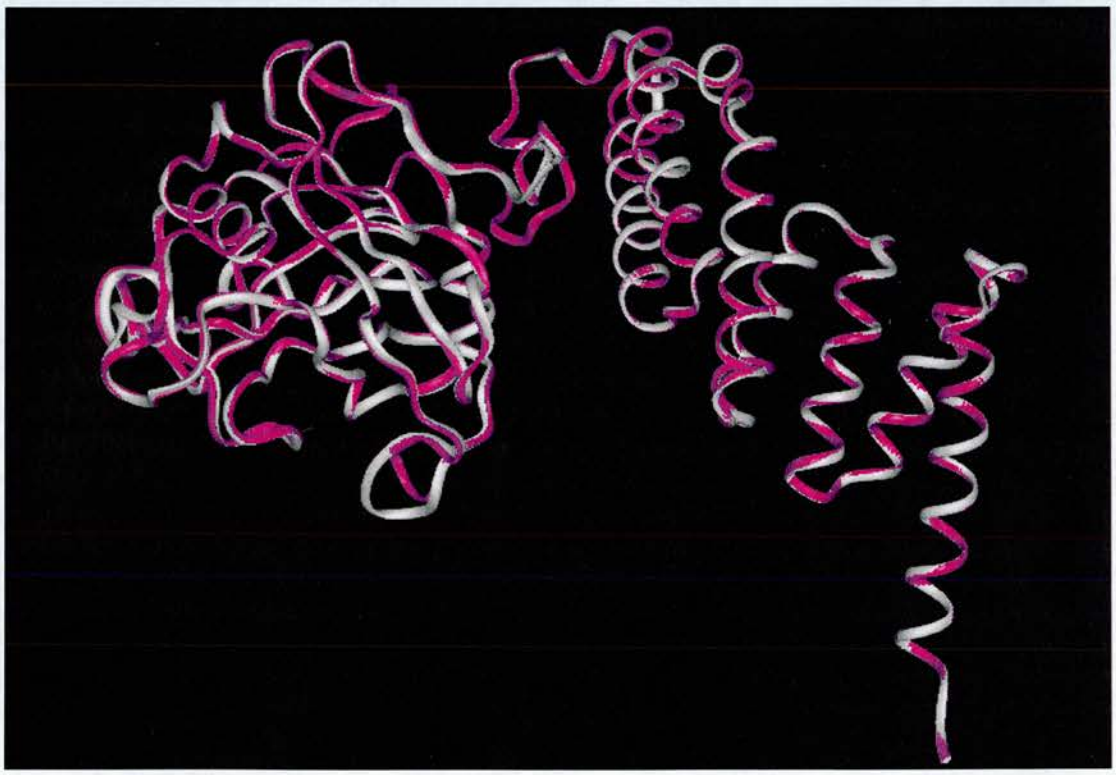


Figure 52
Modelled structure of *L. major* Cyp 40 (white), overlaid on bovine Cyp 40

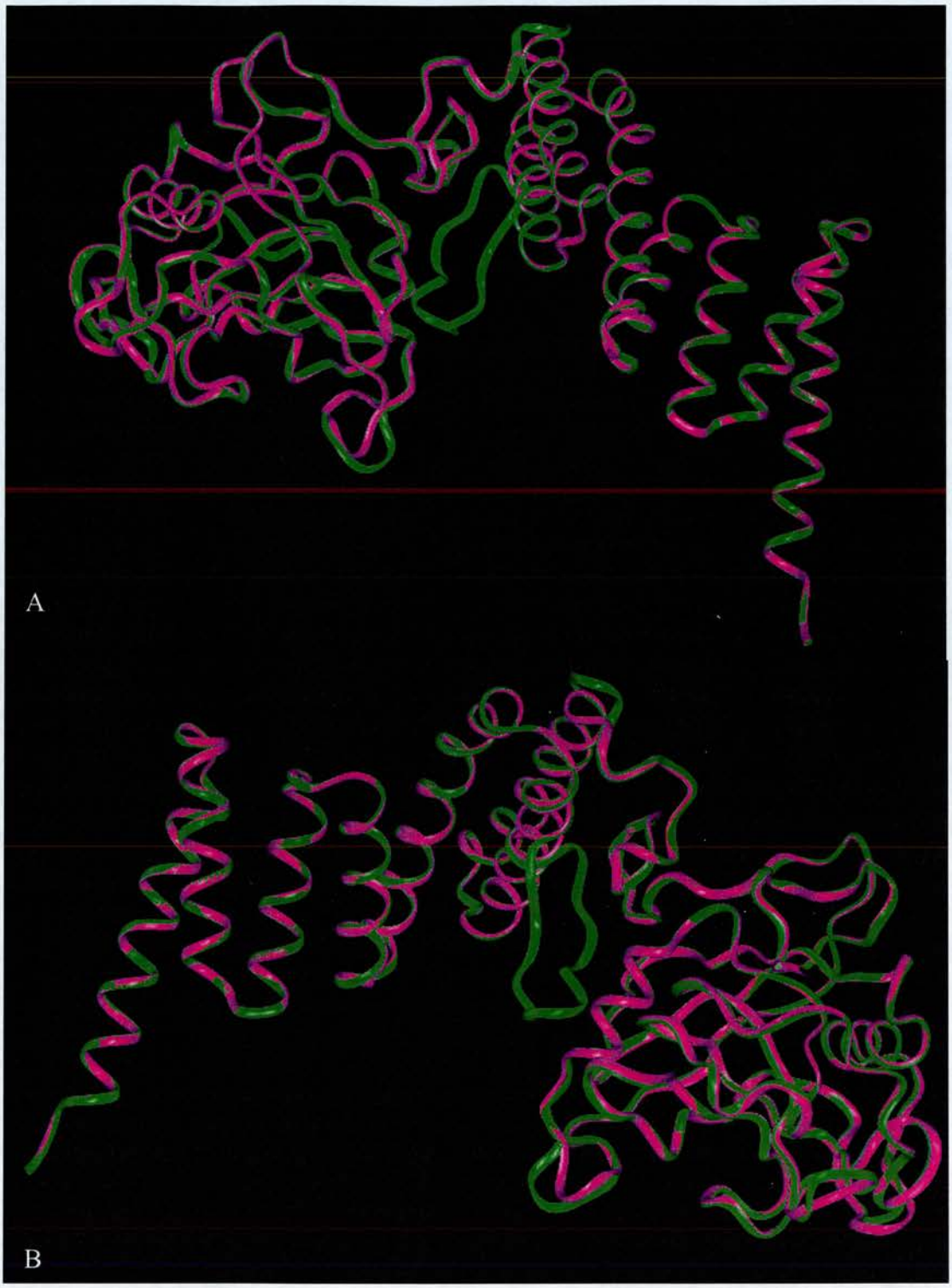


Figure 53

Modelled structure of *D. melanogaster* Cyp 40 (green) overlaid on bovine Cyp 40. Figure A, inner binding groove, Figure B outer binding groove. In each case the inserted loop region from the B helix of TPR 1 can be seen.

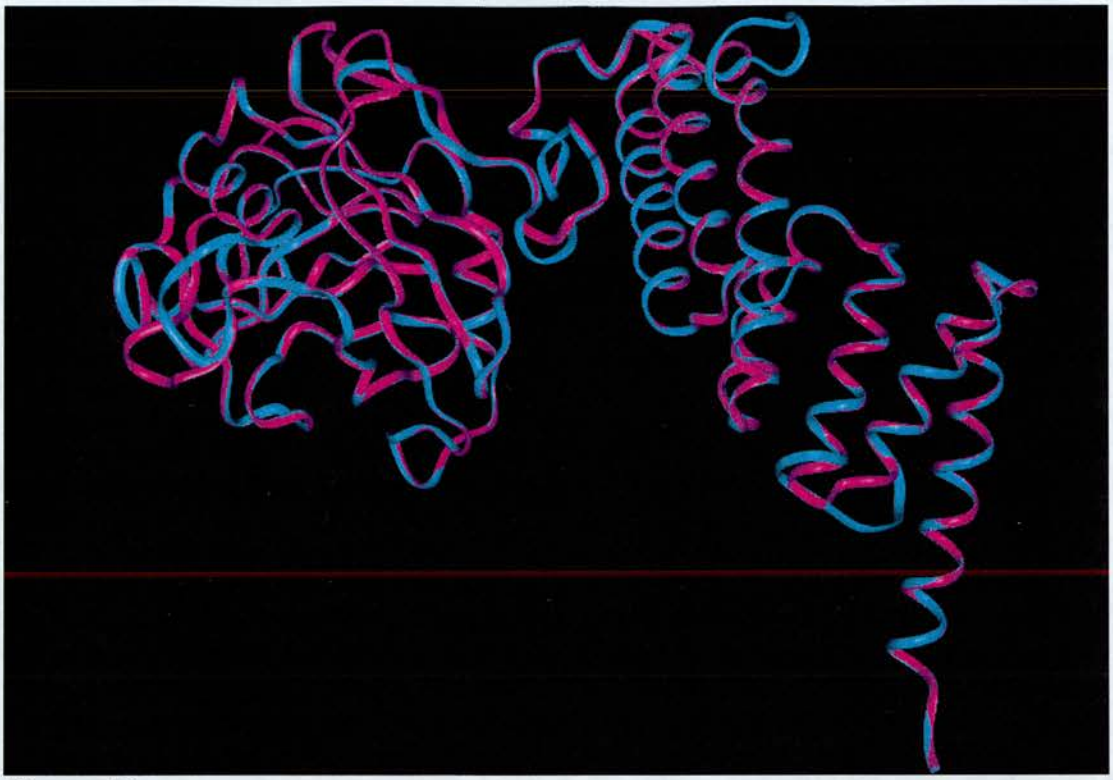


Figure 54
Modelled structure of *S. pombe* Wis2 Cyp 40 (blue) overlaid onto bovine Cyp 40.

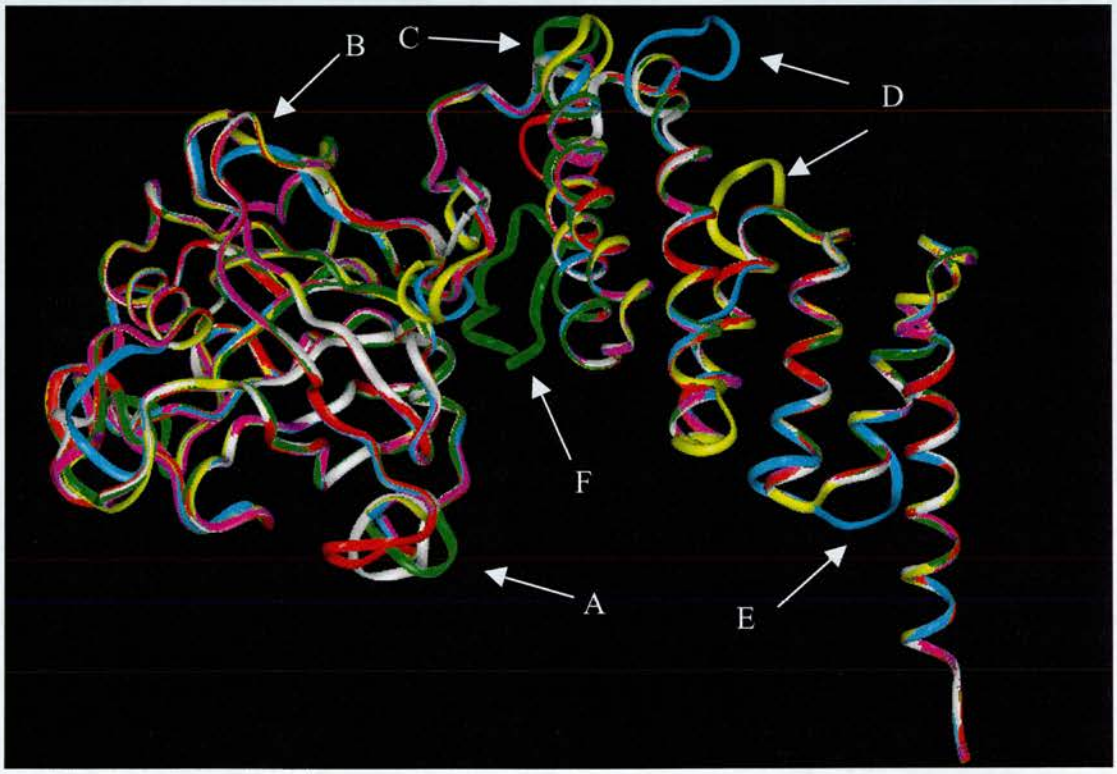


Figure 55
Overlay of all 5 modelled structures (*A. thaliana*, red, *D. melanogaster*, green, *L. major*, white, *N. crassa*, yellow and *S. pombe,* blue) with the experimentally derived structure of bovine Cyp 40 (magenta). Arrows point to areas of structural divergence between modelled and derived structures.

6.4.4

The cyclophilin domain of the modelled proteins is largely conserved between structures

The majority of structurally significant changes within the cyclophilin domain of the modelled structures lie within two surface loops. *Arabidopsis*, *Drosophila* and *Leishmania* diverge from bovine Cyp 40 in the surface loop joining the α-2 helix to the β-8 loop (region A indicated in Figure 55 above). Differences in conformation of this loop have been noted previously, with the suggestion such charged and accessible loops may provide recognition specificity for as yet undetermined partner proteins (Taylor et al., 1998). *S. pombe* Wis2 diverges in the region of the "divergent loop", (region B) largely due to the octapeptide insert in the sequence of Cyp 40 not being present in Wis2. Minor changes are also seen in the linker region of the modelled structures of *L. major*, *N. crassa* and *S. pombe*.

The TPR domain shows significant changes in surface loops and turns

Changes within the TPR domain occur mainly at the ends of helices (regions C, D and E), maintaining the overall super-helical structure of the molecule whilst also potentially allowing specificity of interaction with partner proteins through exposed surface loops and turns. TPR 1 diverges in *A. thaliana*, *D. melanogaster*, *L. major, N. crassa and S. pombe*, TPR 2 in *N. crassa* and *S. pombe* and TPR 3 in *S. pombe*. The most striking area of divergence is seen in the B helix of TPR 1 from *D. melanogaster*.

Drosophila melanogaster Cyp 40 has an inserted loop highly reminiscent of that seen in the structure of $p67^{phox}$.

Figure 53 A and B illustrate two orientations of the modelled structure of *D. melanogaster* Cyp 40. Figure 53 A illustrates the "inner" binding groove and is the orientation used throughout this series of Figures. Orientation B is at 180° to that of A, showing the "outer" face of the molecule. An insertion of twelve amino acids in the sequence was modelled as a loop inserted into the B helix of TPR 1. Secondary structure rendering of the model suggests the inserted loop forms two short regions

of anti-parrallel β-sheets, breaking the B helix of TPR 1 into two shorter helical regions (Figure 56). A similar loop was seen in the structure of $p67^{phox}$ (Lapouge et al., 2000; Grizot et al., 2001). This β hairpin extension in $p67^{phox}$ is inserted between the third and fourth TPR repeats and suggests the modelled loop from *D. melanogaster* may not have been optimally positioned by the automated sequence alignment generated by the modelling programme.

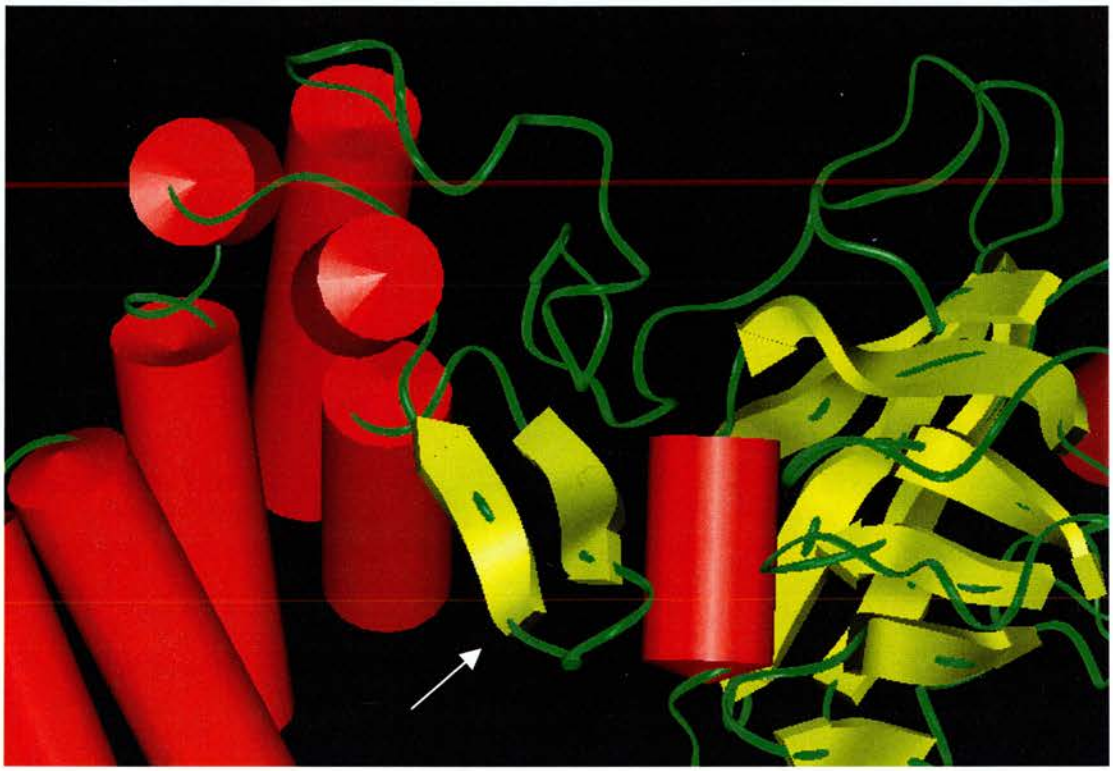


Figure 56

Inserted residues in the sequence of *D. melanogaster* Cyp 40 form a loop containing two anti-parallel β-strands linked by a tight turn. Two residues from this insertion are in close proximity to residues within the cyclophilin domain (within 3Å). The figure was prepared using INSIGHT.

CHAPTER 6B

MODELLED INTERACTIONS BETWEEN THE C-TERMINAL DOMAINS OF HSP90 AND CYP 40.

6.5
INTRODUCTION

Interactions between Hsp 90 and partner proteins

A major biological role for the large immunophilins is binding and regulation of Hsp90 (Frydman and Hohfeld, 1997). Site directed mutagenesis of the serine/threonine phosphatase PP5 identified Lys97, Arg101 and Lys32 to be absolutely required for Hsp90 binding while modification of Arg74 diminished but did not abrogate binding (Russell et al., 1999). A sequence alignment of TPR sequences (Figure 31, Chapter 4) shows that the functionally important residues, Lys97 and Arg101, are conserved at position 2 and 6 of TPR3 for all the related Hsp90-binding proteins.

A comparison of sequences for Hsp90 from different species shows the strongly conserved C-terminal sequence of MEEVD. Deletion mutagenesis studies of the C-terminal domain of Hsp90 have shown that the C-terminal sequence EEVD is important for immunophilin binding (Carrello et al., 1999; Chen et al., 1998).

6.6

ANALYSIS OF MONOCLINIC CYCLOPHILIN 40 CRYSTAL GROWTH AND PACKING PROVIDES INSIGHTS INTO A POSSIBLE INTERACTION WITH HSP 90

6.6.1

Observations from crystal growth and kinetics

An interesting observation made during crystallisation studies of the monoclinic crystal form of Cyp 40 was the exceptionally fast growth along the needle axis (needles grow to over 0.5mm within hours), suggestive of a particularly energetically favourable interaction (Dornan et al., 1999b).

Prior to solution of the structure, a secondary structure prediction of the Cyp 40 sequence gave a strongly predicted C-terminal helix of 28 amino acids in length. This corresponded to a length of 42 Å and is close to the conserved b-cell dimension of the monoclinic needle form of Cyp 40. The b-axis also corresponds to the long and fast-growing needle axis and it is therefore reasonable to suggest that it is this calmodulin-binding helix which determines the length of the b-axis and also the growth kinetics and morphology of monoclinic needle crystals.

6.6.2

Analysis of monoclinic Cyp 40 crystal packing

Analysis of the crystal packing (Figure 57) shows that the only intermolecular contact in the crystal occurs upon stacking the molecules along the c-axis with the C-terminal residues EKAAY fitting in tightly to a groove in the surface formed by the three TPR motifs. The residues defining the groove are Lys227, Phe234, Ser274, Asn278, Lys308 and Arg312. Significantly, the interaction involves close contacts with the three positively charged residues identified in the site directed mutagenesis study as being important for interaction with Hsp90 (Russell et al., 1999).

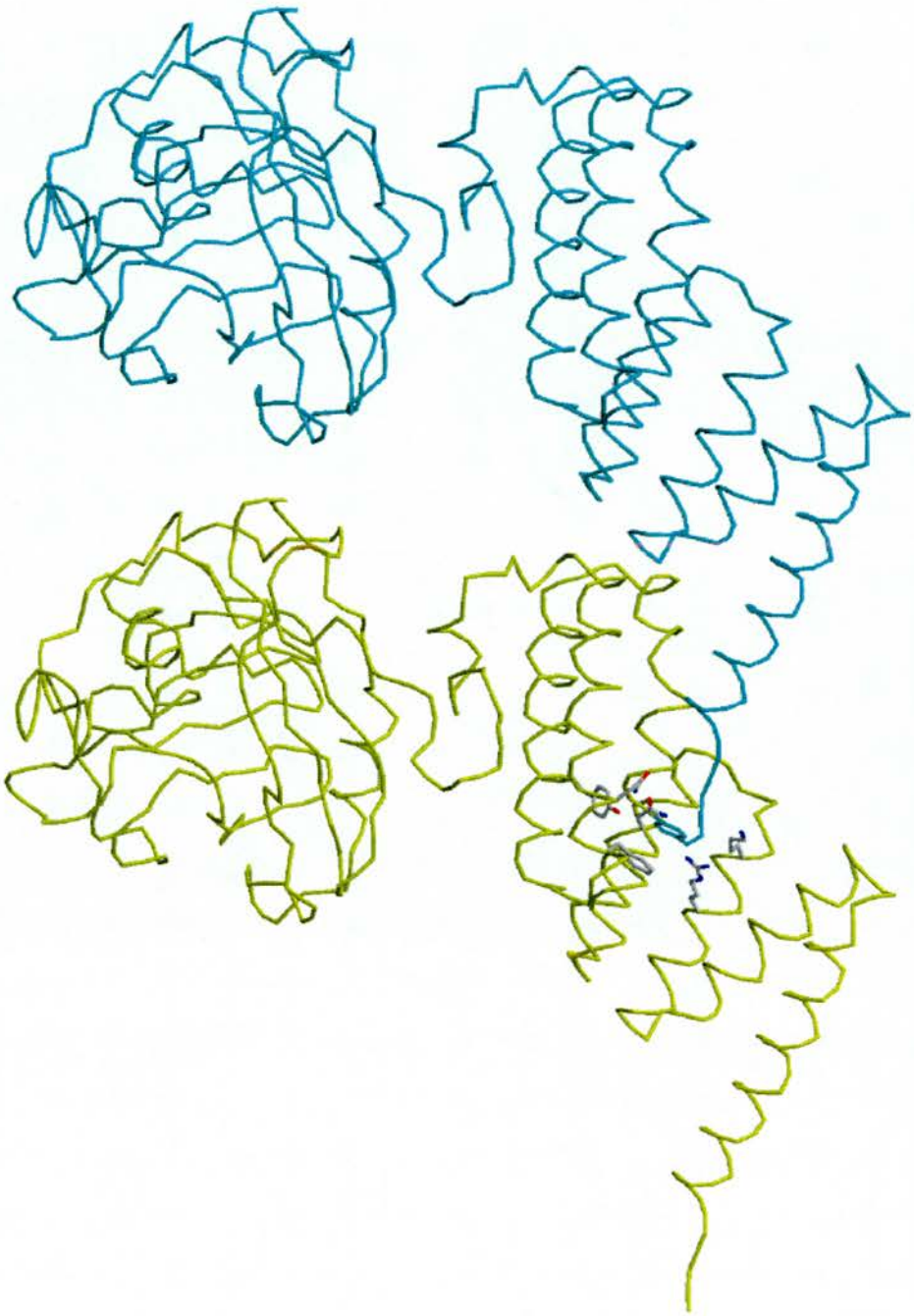


Figure 57

Picture of the main intermolecular interaction in the monoclinic form of Cyp40.

The yellow and blue molecules are related by a unit cell translation along b of 47.3 Å. The C-terminal residues EKAAY one molecule (blue) fit into a groove of an adjacent molecule (yellow). The groove is formed by two residues from each of the three TPR helices ($A_{1,2,3}$): Lys227, Phe234, Ser274, Asn278, Lys308 and Arg312 (side chains shown). Tyr 365 fits into a hydrophobic pocket in the groove and is the last residue visible in the electron density (The C-terminal five residues are presumably structurally disordered in the crystal).

6.7

MODELLING THE INTERACTION BETWEEN THE HSP 90 C-TERMINAL PENTAPEPTIDE MEEVD AND THE TPR REPEAT DOMAIN OF CYP 40.

6.7.1

The X-ray structure of the Cyp40 TPR domain provides a model binding surface for Hsp90.

6.7.1.1

Modelling the Cyp 40/MEEVD peptide interaction

The intermolecular interaction from crystal packing (Figure 57) was used as a guide to model the Cyp40 interaction with the Hsp90 C-terminal MEEVD sequence.

The structure of Cyp 40 was unaltered throughout the modelling exercise. The MEEVD pentapeptide was manually docked into the binding groove, mimicking the interactions of EKAAY with each of the A helices of the TPR motif. Manual docking provided a sterically reasonable fit such that each of the four carboxyl groups of the MEEVD sequence were at distances appropriate to the formation of a salt bridge with Lys308, Arg312 or Lys227. Furthermore; the hydrophobic pocket occupied by Tyr365 in the crystal structure was occupied by valine. The methionine also filled an available hydrophobic pocket.

All modelling was carried out using WITNOTP (Widmer, 1997).

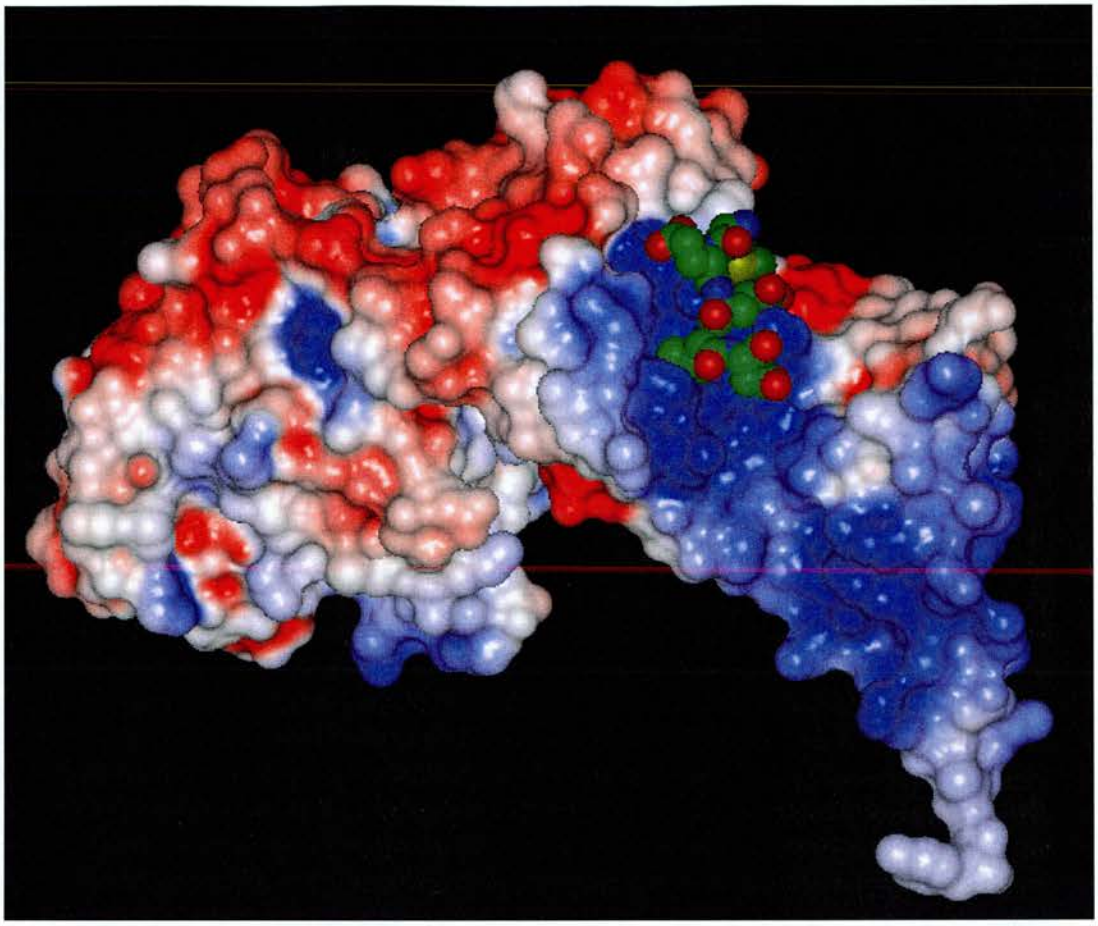


Figure 58

Picture of modelled MEEVD docked onto the Hsp 90 binding groove. An electrostatic surface was calculated for Cyp 40 using the DelPhi module from the INSIGHT II suite of programs. MEEVD was rendered in space filling CPK. This representation of the two molecules clearly shows MEEVD docking into the concave inner groove, contacting residues from each of the A helices of the TPR motif.

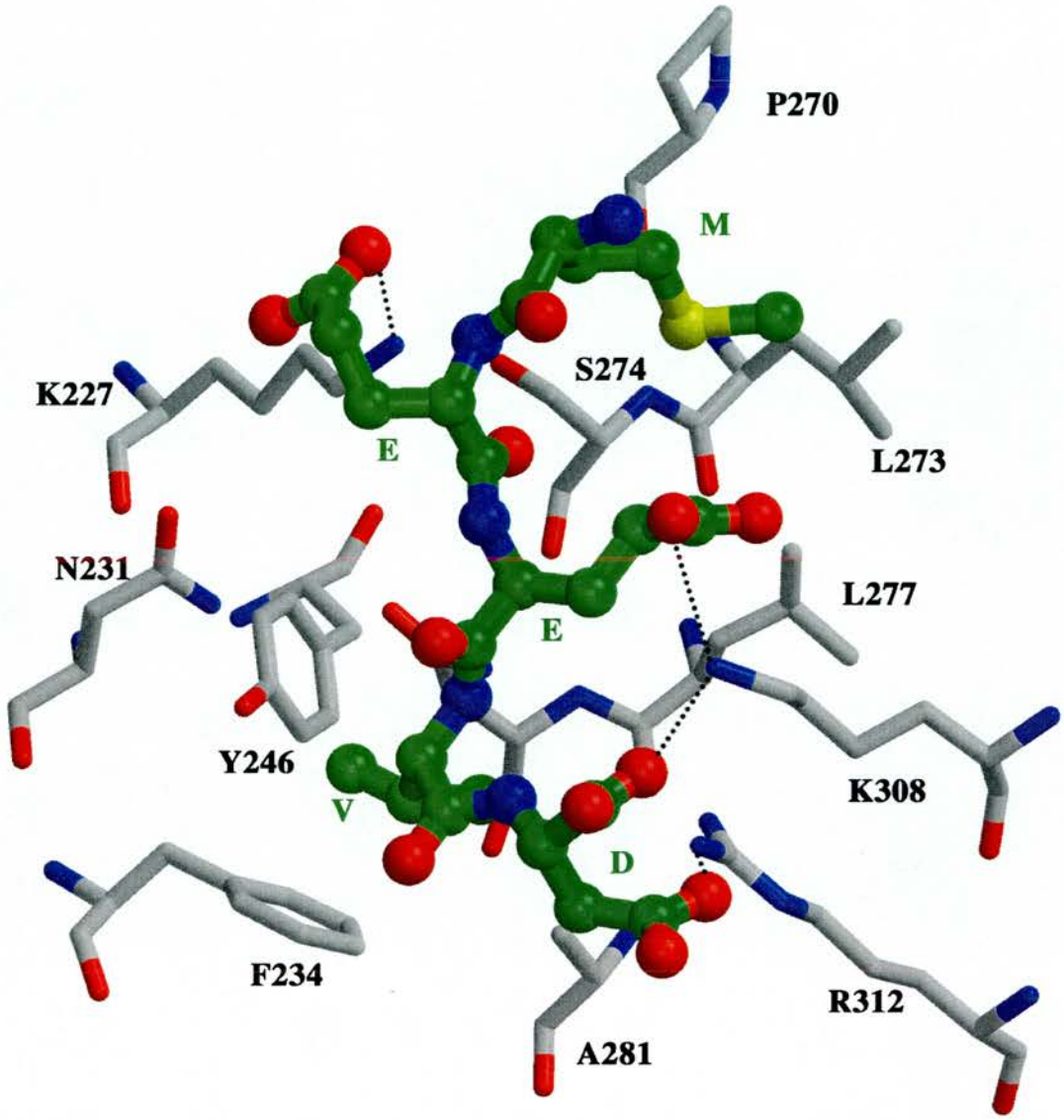


Figure 59

"High resolution" picture of the conserved C-terminal Hsp90 sequence (MEEVD) docked into the putative binding groove of Cyp 40. The MEEVD C-terminal sequence of Hsp 90 is shown in green as balls and sticks. Those residues defining the putative binding groove in the Cyp40 structure are coloured grey. The conformation of the MEEVD peptide was selected to mimic the interaction shown in Figure 57. The Cyp40 structure was not altered. Conformational angles for MEEVD backbone and side chains all lie within acceptable values. All negatively charged groups can form salt bridges with K227, K308 and R312. The valine and methionine side chains sit deep in hydrophobic pockets.

6.7.1.2

Two possible Hsp 90 interaction modes revealed by MEEVD binding

The modelled interaction between the MEEVD motif of Hsp90 and Cyp 40 provides an alternative to that seen in the X-ray structure of the complex between the TPR-2A domain of Hop with the Hsp90 MEEVD motif (Scheufler et al., 2000). An overlay of the two structures is shown in Figure 60.

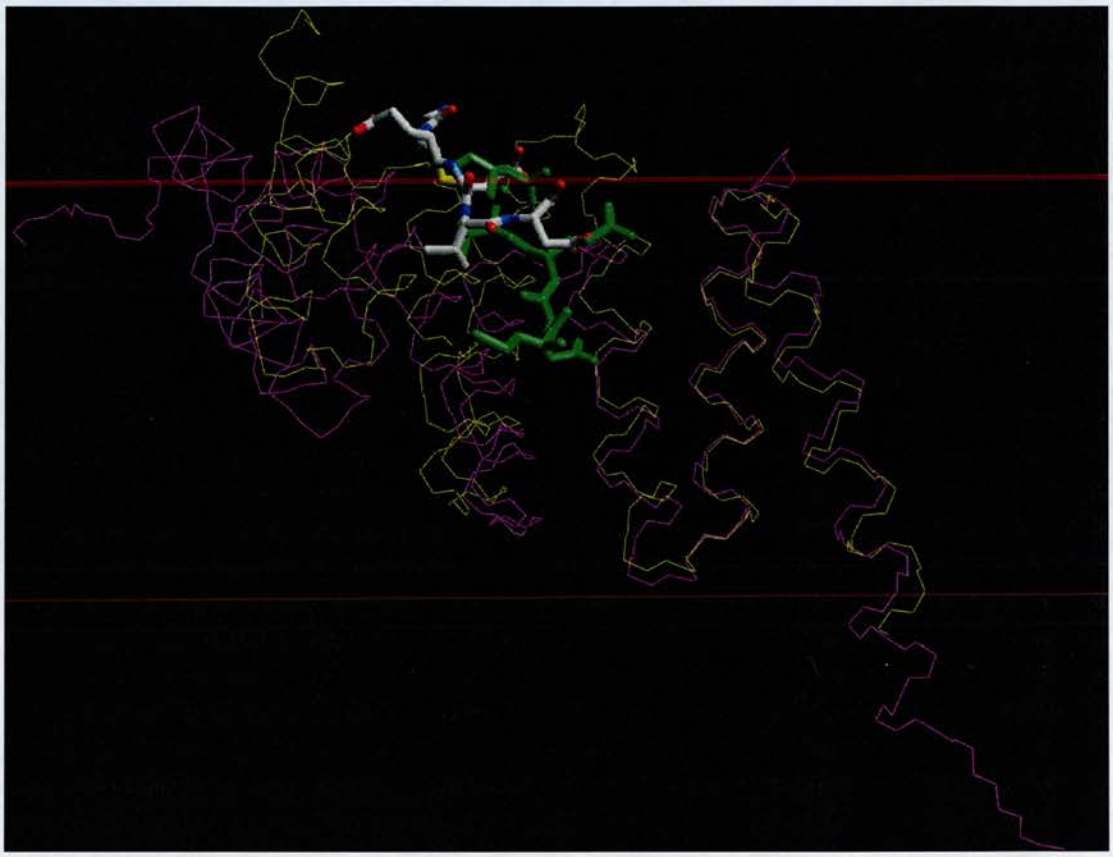


Figure 60
View of an overlay of the X-ray structures of Cyp 40 and the TPR-2A domain of Hop, solved as a complex with MEEVD. The main chain trace of Cyp 40 is shown in magenta, residues 223 to 352 of human Hop are shown in yellow. The modelled MEEVD is shown as capped sticks, coloured by atom type. MEEVD from the structure in complex with the TPR-2A domain of Hop is shown in green capped sticks.

This model requires the MEEVD peptide to run in the opposite direction to that found in the Hop-peptide complex (Scheufler et al., 2000) suggesting the possibility that the binding groove of TPR domains may act as a peptide binding site with broad specificity.

6.7.2

Interaction between the 95 C-terminal residues of Hsp 90 with the TPR domain of Cyp 40

6.7.2.1

Docked MEEVD provided a starting point for model development

To further characterise the proposed model of C-terminal regions of Hsp 90 interacting with Cyp 40, the peptide model was expanded to include the 90 residues preceding MEEVD in the protein sequence.

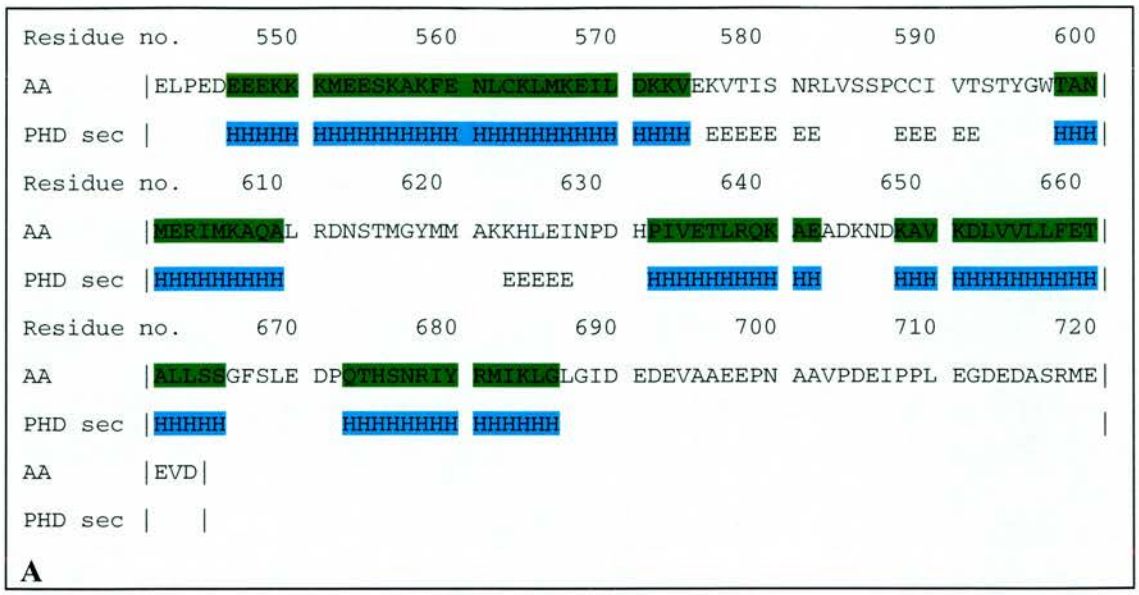
Secondary structure predictions of full length Hsp 90 suggest the C-terminal domain is highly helical, with an unstructured tail region. The interaction domain with Cyp 40 resides within residues 600-724, as identified by deletion mutagenesis interaction studies. Development of a model incorporating those residues, taking into account predictions of secondary structure elements would therefore appear to be a sensible starting point.

6.7.2.2

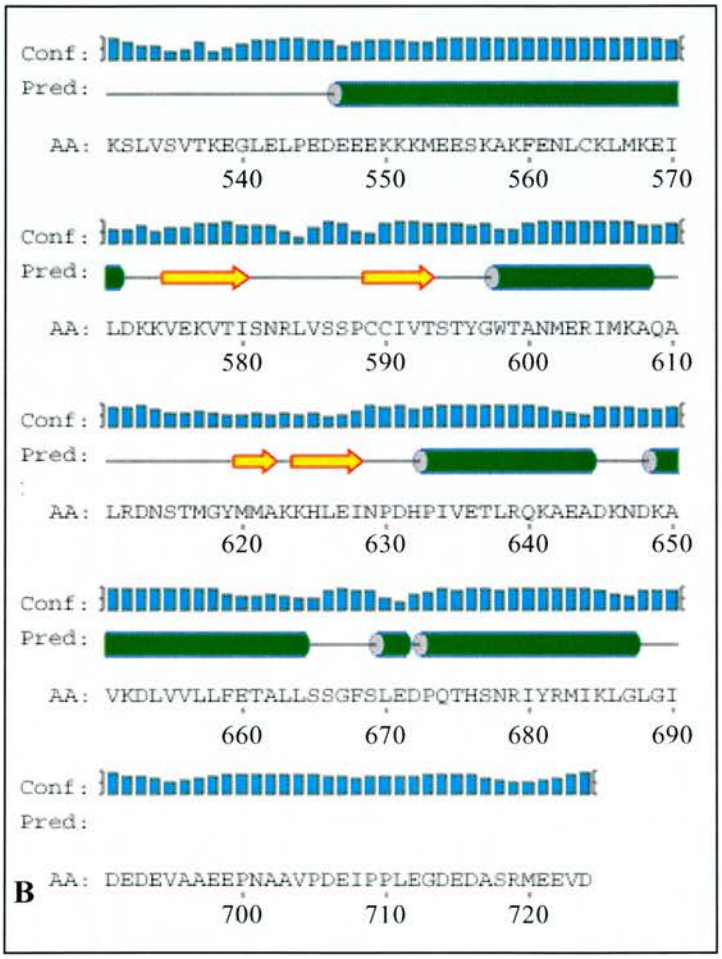
Secondary structure prediction

Secondary structure predictions of full length Hsp 90β were calculated by two independent programmes, PredictProtein (Rost et al., 1994), (http://cubic.bioc.columbia.edu/predictprotein) and PSIpred (McGuffin et al., 2000), (http://insulin.brunel.ac.uk/psipred/). The resulting predictions were very similar, with 5 helices strongly predicted in the sequence from residue numbers 540 to 724. The modelled MEEVD binding site provided one anchor point for the C-terminal region, incorporation of additional residues required the model to complement the original proposed interaction.

Figure 61 shows the results obtained from the PSIpred, and PredictProtein servers.



A



B

Figure 61
Graphical representation of the output from both prediction of secondary structure websites. Panel A illustrates the results from PredictProtein and shows only predicted secondary structure elements, all other data was edited out prior to incorporation into the figure. Helices are represented by the letter H in cyan, and highlighted in the sequence. in green. Panel B, results from PSIpred. Helical regions are delineated by green cylinders, sheets by yellow arrows. The boxed regions in cyan illustrate the level of confidence in the prediction for individual amino acids. The results from both programs are very similar.

6.8

MODELLING THE INTERACTION BETWEEN CYP 40 AND THE FINAL THREE C-TERMINAL HELICES OF HSP 90β

6.8.1

Overview of modelling strategy

The aim of this part of the project was to utilise insights gained from the structure of Cyp 40 to define a binding site for the C-terminal region of Hsp 90. The working hypothesis developed was of helical Hsp 90 docking into the groove formed by the TPR domain of Cyp 40, with the terminal pentapeptide docked as illustrated above.

Secondary structure predictions suggested that a three or four-helix bundle would be an appropriate template model for the C-terminal domain of Hsp90 (residues 630 – 724). This structural motif was chosen as it allowed modelling with a topologically appropriate structure, for the following reasons:

- Basing the modelled structure on a three or four-helix bundle ensures a spatially and topologically accurate model for the three helices of Hsp 90
- Alignment of the first helix of the bundle with the C-terminal long helix of Cyp 40 would maintain the correct topology for the molecule
- Maintenance of integrity of the bundle when aligned with Cyp 40 ensures a sterically reasonable "fit" between the two molecules, the remaining helices dock into the Cyp 40 binding groove
- The resultant model incorporates the required elements, a bundle of three helices, accommodated within the Cyp 40 binding groove.

6.8.2

Selection of an appropriate template model

The CATH protein structure classification database

The CATH protein structure classification database was searched for an appropriate template molecule. This database uses protein structures, solved to a resolution of at

least 3Å and deposited with the Brookhaven Protein Databank. Structures are arranged in a hierarchical manner, using the following four main features to assign each structure:

- **Class**, 3 main divisions (mainly alpha, mainly beta, alpha/beta)
- **Architecture**, determined by overall shape of the domain structure
- **Topology (fold family)**, determined by the overall shape and connectivity of the secondary structure elements
- **Homologous superfamily groups,** protein domains thought to share a common ancestor.

Restraints imposed upon the model

The "interaction groove" in the structure of Cyp40 defined the topological requirements of the model, many structures from the database could be discarded by simple visual inspection of molecular architecture. A number of potentially suitable template molecules were chosen for closer examination, these included two structures of ligand binding domains from the aspartate receptors of *E. coli* and *S. typhimurium* (PDB ID 2ASR and 1LIH respectively). The PDB files for selected structures were exported from the database and displayed using the programme WITNOTP (Widmer, 1997). Template structures were aligned with and anchored along the long C-terminal helix of Cyp 40, running N to C terminally. The model structures were then examined to determine that which appeared to have the best fit, when compared to the arrangement of secondary structural elements predicted for Hsp90. The ligand binding domain of the aspartate receptor from *E.coli* (PDB ID no. 2ASR) was chosen as the best available model structure.

The first helix of the 4 helix-bundle was anchored to the long helix of Cyp 40, subsequent helices were defined by their spatial relationship to this anchoring helix. Key structural elements from the Hsp90 secondary structure predictions ie positions of turns, lengths of helices and in particular the location of the helix-turn-helix-motif (residue number Lys 649 to Gly 687 in the sequence) were mapped onto the template structure. Positioning the helix-turn-helix motif was of particular importance as residues within this region had been shown by site directed mutagenesis studies and

bio-affinity interactions to be crucial in the interaction with the tetratricopeptide repeat domain of Cyp 40 (Carrello et al., 1999) .

6.8.3

Preliminary model of three Hsp 90 C-terminal helices docked onto the surface of Cyp 40

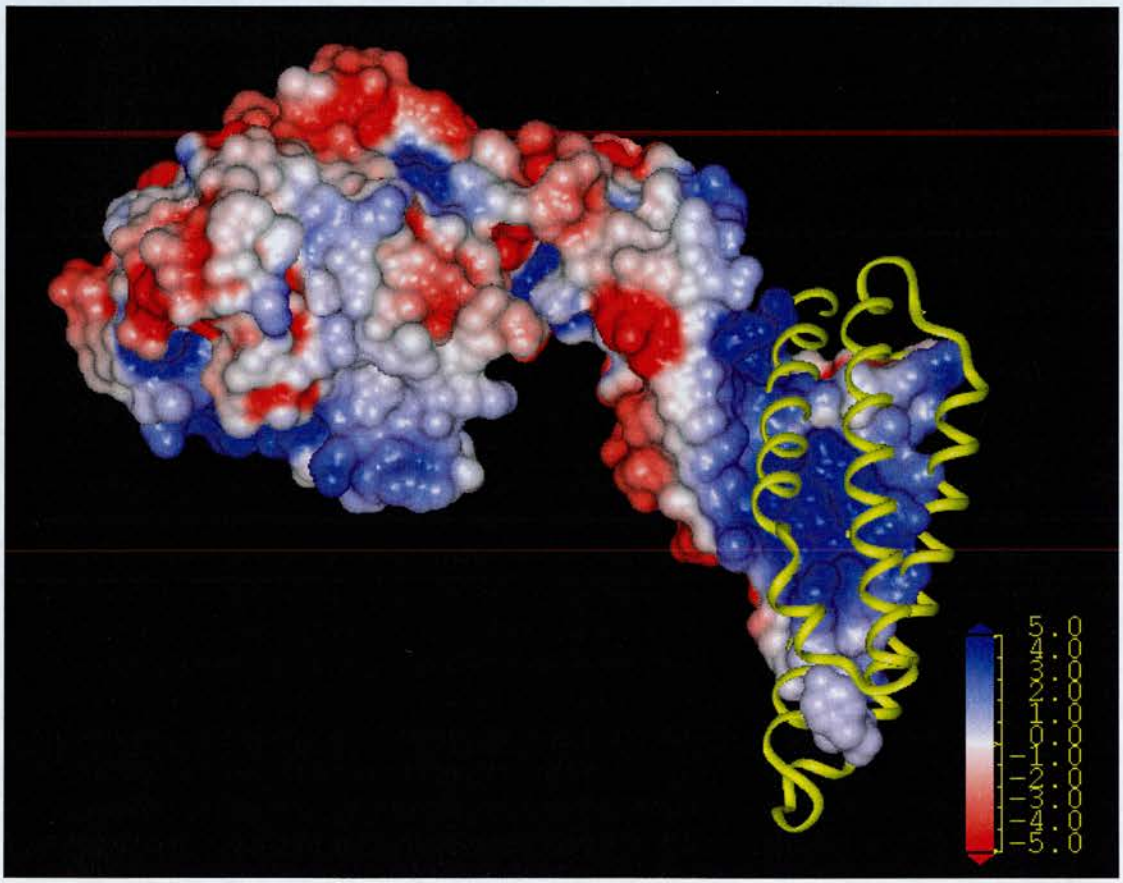


Figure 62

The full length structure of the ligand binding domain of the *E. coli* aspartate receptor (PDB ID 2ASR) was used for preliminary modelling studies. The structure was manually docked into the Cyp 40 binding groove, the N-terminal helix of 2ASR aligned along the C-terminal putative calmodulin binding helix of Cyp 40. PDB files of the two structures docked in WITNOTP were used to prepare the above figure in INSIGHT, relative positions of the two proteins were unaltered following manual docking.

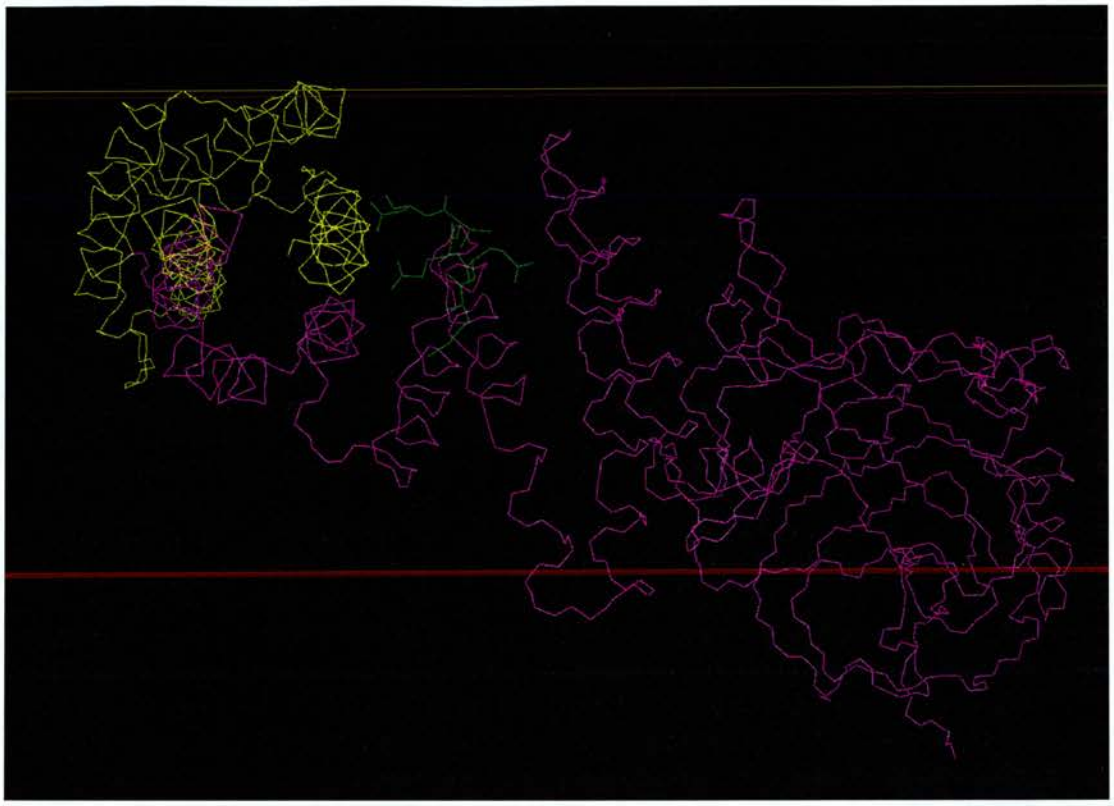


Figure 63

Representation of the modelled interaction between the ligand binding domain of the *E. coli* aspartate receptor (yellow) and Cyp 40 (magenta). The first helix from the receptor can be seen aligned along the C-terminal calmodulin binding helix of Cyp 40, the remaining helices occupy the inner binding groove of Cyp 40. Manually docked MEEVD is shown in green. The figure was prepared using WITNOTP.

To help in the visualisation of the interaction between the helices of Hsp90 and the interaction groove of Cyp 40, regions from the helices of the aspartate receptor were cut out from the molecule to represent lengths of helices predicted for Hsp90. These selected residues from the sequence were mutated to correspond to those in the appropriate predicted helical regions of the Hsp90 molecule. The mutated residues were also renumbered as appropriate for the Hsp90 sequence.

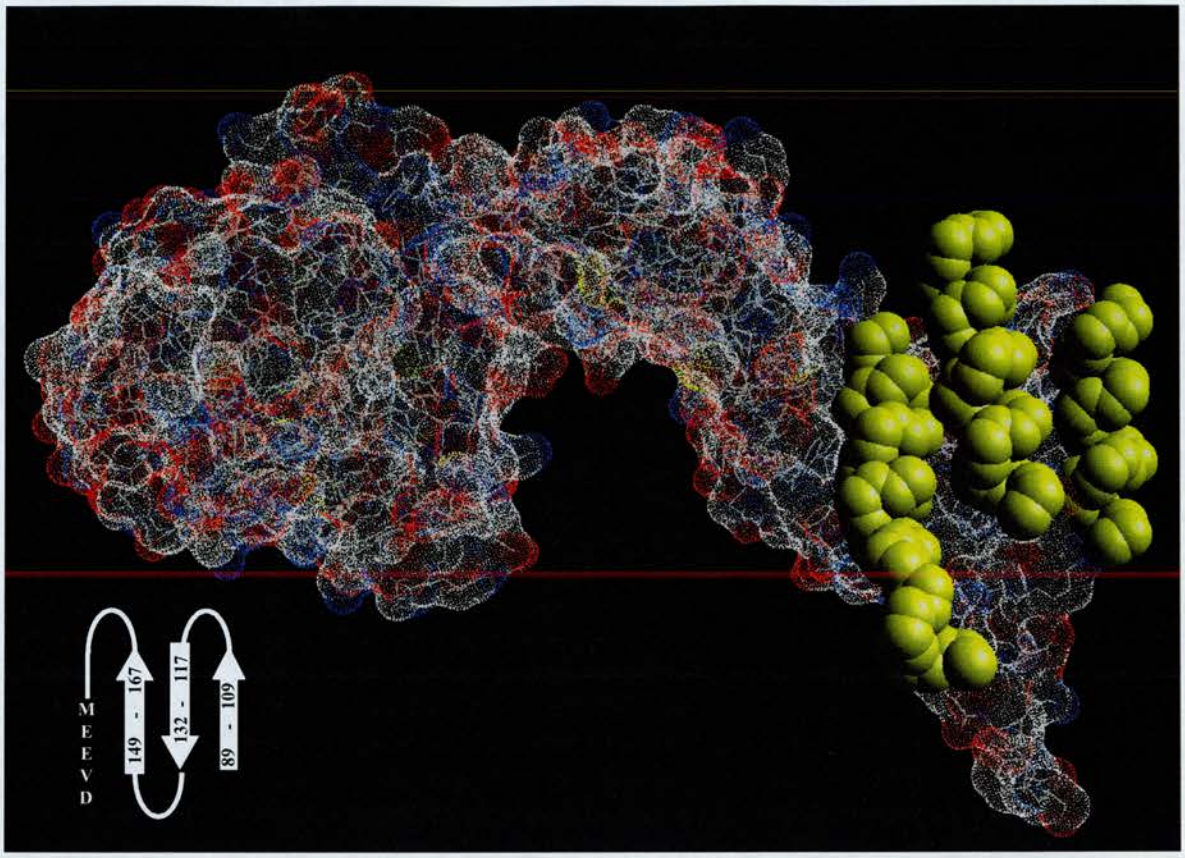


Figure 64

Modelled interaction of 3 predicted Hsp90 helices docking into the Cyp 40 binding groove. The Cyp 40 molecule is displayed with a Connolly surface, the helical regions of Hsp90 are displayed in space filling CPK. The schematic diagram illustrates the orientation of each of the helices with respect to the Cyp 40 structure and the MEEVD pentapeptide.

Binding studies suggests that the TPR domain alone is not sufficient for the stable association of these immunophilins with Hsp90, but requires additional contributions from flanking sequences upstream and downstream of the TPR region (Ratajczak and Carrello, 1996). The proposed C-terminal Hsp90 interaction with Cyp 40 must therefore be regarded as a model of one of a number of possible immunophilin-Hsp90 interactions. Due to the high degree of conservation of the EEVD motif, the interaction might possibly be driven by the initial recognition of and docking with this particular motif to regions in partner proteins, followed by stabilising interactions with other residues structurally distant from the MEEVD/Cyp 40 interface.

6.9
DISCUSSION

Modelled structures of five Cyp 40 homologues. The cyclophilin domain is very structurally conserved between the different species represented, as may have been expected following analysis of previously published structures.

Maintenance of the overall super-helical fold of the TPR domain from these proteins is interesting, suggestive of the necessity of maintaining this putative interaction domain (the binding groove), whilst allowing specificity of interaction with partner proteins through variable regions principally located in surface loops and turns between individual TPR motifs. *D. melanogaster* Cyp 40 is the most divergent structurally, containing a large insertion forming two antiparallel β sheets, located in the model in the B helix of TPR1. Analysis of the model suggests that whilst structurally this is a reasonable model for the region, it is probably inappropriately located within a TPR motif and would be better placed at the end of the B helix of TPR1.

Modelling the interaction of the C-terminal domain of Hsp90 with Cyp 40 resulted in a hypothesis in which MEEVD docks into the binding groove of Cyp 40, which also accommodates three strongly predicted Hsp90 helices. Mutagenesis studies suggest other regions in these molecules are also important to stabilise and maintain this interaction (Ratajczak and Carrello, 1996). Site point mutational analysis of both the Cyp 40 TPR domain and the C-terminal region of Hsp90 would provide further insight into residues crucial for the interactions between these two proteins. The model presented above might be useful in helping to select the most structurally appropriate residues to target for such studies.

Interactions between immunophilins and Hsp90 are without doubt complex, the model presented above must therefore be considered as one of a number of possible modes of interaction.

CHAPTER 7.

CRYSTALLISATION AND PRELIMINARY PARTIAL X-RAY STRUCTURE DETERMINATION OF FKBP 22 FROM NEUROSPORA CRASSA

7.1
INTRODUCTION

This chapter will focus on describing characterisation and structural studies carried out on a recently described member of the FKBP family, FKBP 22 from the fungus *Neurospora crassa* (Solscheid and Tropschug, 2000). Recently isolated from the endoplasmic reticulum, the N-terminal domain of FKBP 22 was shown to be homologous with that of the well characterised FKBP 12. Sequence analysis shows the peptidyl-prolyl active site to be highly conserved; measured enzymatic activity is similar to that for other members of the family. In common with the smaller FKBP proteins, FKBP 22 has been shown to bind FK506 and rapamycin with IC_{50s} in the nM range.

The protein is synthesised as a precursor protein, with 20 additional amino acids at the amino terminus, typical of an ER targeting signal. The mature protein, as calculated from the cDNA sequence, has a molecular weight of 20,912 Da. Electrospray mass spectrometry of the purified protein however yielded a mass of 22,043 Da, therefore the authors called the protein FKBP 22. The extreme C-terminus of the protein contains a putative ER retention signal (-HNEL).

The N-terminal domain of NcFKBP22

The N-terminal domain of NcFKBP 22 is a typical FKBP domain, with highest homology to the cytosolic/mitochondrial NcFKBP13, (60%), with other % identities averaging 50%. All residues important for PPIase activity and FK-506/rapamycin binding are present in the FKBP domain of NcFKBP22.

The C-terminal domain of FKBP 22

The C-terminal domain of the protein (residues 116-197) is unlike any other FKBP that has been characterised to date. The extreme carboxy-terminus has a biased composition of amino acids, with a clear bias against Phe, Gly, Ile, Leu, Pro and Arg, whilst Ala, Glu, Lys and Val are over-represented. This domain also has a very acidic nature and is consistent with an amphipathic helical conformation. BLAST searches using the C-terminal domain revealed a low sequence similarity to short regions from late embryogenesis abundant (LEA) proteins from a number of different species. These LEA proteins are thought to protect cells from stress associated with dehydration.

The extreme C-terminal sequence (HNEL) is an ER retention signal; it has been suggested that NcFKBP22 may be involved in protein folding in the ER, with the charged C-terminal domain acting as a recognition site for target proteins.

7.2
MATERIALS AND METHODS

7.2.1
Isolation and cloning of full–length cDNA encoding NcFKBP 22

Members of the Tropschug group, University of Freiburg, Germany, carried out all cloning of NcFKBP22.

FKBP 22 was first isolated from a preparation of *N. crassa* hyphae. The hyphal lysate was applied to an affinity chromatography resin that comprised the FK-506 homologue ascomycin (Kay, 1996) coupled to Sepharose beads. After extensive washes, specifically bound proteins were eluted from the resin by a shift in the pH from 7.4 to 2.5, SDS-PAGE analysis of eluted proteins followed. In addition to NcFKBP13 previously isolated from *N. crassa* (Tropschug et al., 1990), another protein with an apparent molecular weight of 29kDa was also routinely seen from these preparations.

This affinity-purified material was N-terminally sequenced and shown to share 12 out of the 33 identified amino acids with NcFKBP13. A full length cDNA clone of NcFKBP22 was prepared, allowing deduction of the full amino acid sequence of this novel protein.

The following is provided for completeness. A λgt11 library of *N. crassa* was used for screening. One positive clone was identified that contained a short cDNA insert, encoding for a C-terminal part of NcFKBP22. Reverse primers from this region were used in combination with degenerate primers directed against the N-terminus to isolate a specific NcFKBP PCR product that was finally used in plaque hybridisations to isolate full-length cDNA. After restriction digestion, the fragment was cloned into a suitably digested pQE-32 vector (Quiagen). Ligation and transformation protocols of the Quiagen pQE vector manuals were followed. The plasmid DNA was then transformed into competent cells for subsequent recombinant protein expression.

7.2.2

pET expression of recombinant FKBP 22 in E.coli

The standard protocol developed throughout the course of this project, for large-scale over-expression of proteins under control of the T7 promoter was used (see Chapter 2, Section 2.3) for FKBP 22. All work was carried out either on ice using chilled solutions, in equipment pre-cooled to 4°C or at 4°C in the cold room.
His-tagged FKBP 22 was expressed following a four hour induction with 0.4mM IPTG (Melford Laboratories).

7.2.3

Purification of recombinant FKBP 22

The construct contained an N-terminal 6-His tag to simplify purification of recombinant protein. The method to purify His tagged proteins under native conditions, as outlined in the Qiagen Handbook was followed, and is included below for completeness.
Cell pellets were defrosted on ice and solubilised in lysis buffer containing 50mM NaH_2PO_4, 300mM NaCl, 10mM imidazole at pH 8.0, at a ratio of 2-5ml per gram wet weight of cell pellet. Lysozyme (ICN/FLOW) was added to 0.1% w/v and incubated on ice for 1 hour. An equal volume of deionised water at 4°C was added and the cell suspension mixed. The lysate was centrifuged to remove cell debris; the resulting clarified lysate was frozen at –20°C.

Clarified cell lysate was applied to a 10ml column of Ni-NTA Superflow in buffer A (50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole) at pH 8.0. The column was washed until a baseline was established, followed by elution of the tagged protein by application of a linear gradient of buffer B (50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole) at pH 8.0. Column fractions were tested on 15% SDS-PAGE gels and fractions containing the protein of interest were pooled.

7.3

CHARACTERISATION OF FKBP 22

7.3.1

Dynamic light scattering experiments

Dynamic light scattering experiments on the purified FKBP 22 were carried out using the DynaPro 801 with microsampler attatchment. A number of protein concentrations were tested, the protein was monitored for stability and tendency to aggregate.

7.3.2

Electrospray ionisation mass spectrometry

Purified protein was further characterised by mass spectrometry analysis providing information on the parent ion(s) present. The sample was introduced into the instrument stream via an infusion pump, the m/z spectrum deconvoluted using appropriate software. Thanks to Dr. Andrew Cronshaw for the mass spectrometry results.

7.3.3

Crystallisation of FKBP 22

The hanging drop vapour diffusion method as described in chapter 2, Section 2.5.1.3 was used throughout this series of experiments.

7.3.4

Data collection

Crystals were transferred from the hanging drop using an appropriately sized cryo-loop into a cryo-protectant solution containing mother liquor plus 20% glycerol. Crystals were removed from the cryo-protectant solution after 5 seconds and mounted in a cryo-loop and flash frozen in liquid nitrogen. X-ray data were collected at 100K using a MAR image plate and processed using DENZO (Otwinowski and Minor, 1993).

7.4

PARTIAL STRUCTURE DETERMINATION OF FKBP 22

7.4.1

Structure solution and refinement

Members of the Structural Biology Group, ICMB, University of Edinburgh, determined the partial structure. Professor Malcolm D. Walkinshaw obtained the initial partial structure solution, using molecular replacement; with FKBP 12 as a model. Dr. Joel Tyndall carried out further refinement of this structure.

The partial structure was solved using data to 2.6Å resolution, only the FKBP domain is well defined in density.

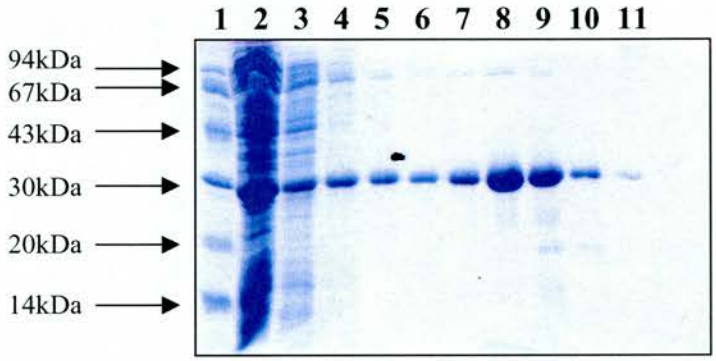
7.5

RESULTS AND DISCUSSION

7.5.1

Purification of recombinant FKBP 22

The FKBP 22 purification illustrated was carried out by Miss Monludee Opamuwutthikul as part of her PhD programme.



A representative gel illustrating a typical elution profile from Ni-NTA resin. Lane 1, molecular weight markers, lane 2, cleared lysate, lanes 3, 4 and 5, wash fractions. Lanes 6 – 11, elution of FKBP 22 with a linear gradient of imidazole. Protein was eluted at approximately 40%, corresponding to 100mM imidazole.

Figure 65

Purification of FKBP 22 was achieved in one affinity chromatography step.

7.5.2

Dynamic light scattering analysis

Table 18 illustrates a typical data-set obtained from a dynamic light scattering experiment carried out using purified FKBP 22. Measurements were taken at room temperature using a DynaPro 801 instrument with micro-sampler attachment.

Measurement number	**Polydispersity (nm)**	**Estimated Molecular weight**	**Baseline**	**SOS Error**
1	0.0	63 kDa	1.001	4.939
2	2.0	69kDa	1.000	6.488
3	1.0	82kDa	1.000	3.529
4	1.3	77kDa	1.003	1.818
5	1.0	58kDa	0.999	4.060
6	1.3	83kDa	1.000	1.515
7	0.0	82kDa	1.001	3.711
8	2.0	88kDa	1.002	3.466
9	0.0	87kDa	1.004	1.342
10	2.5	91kDa	1.001	3.008
11	1.7	119kDa	1.017	2.485
12	2.3	84kDa	1.000	3.488
13	0.0	70kDa	0.998	2.746
14	1.8	64kDa	0.999	2.619
Mean values	**1.2**	**80kDa**	**1.002**	**3.230**

Table 18

Purified FKBP 22 at 6mg/ml was analysed as described in Chapter 2, utilising the room temperature micro-sampler attachment of the DynaPro 801. Bimodal analysis of the data suggests 96% of the protein has an estimated molecular weight of 29kDa, the remainder of the sample consisting of a large aggregate. These results also suggest the possibility the protein may be dimeric.

7.5.3

Mass spectrometric analysis of FKBP 22

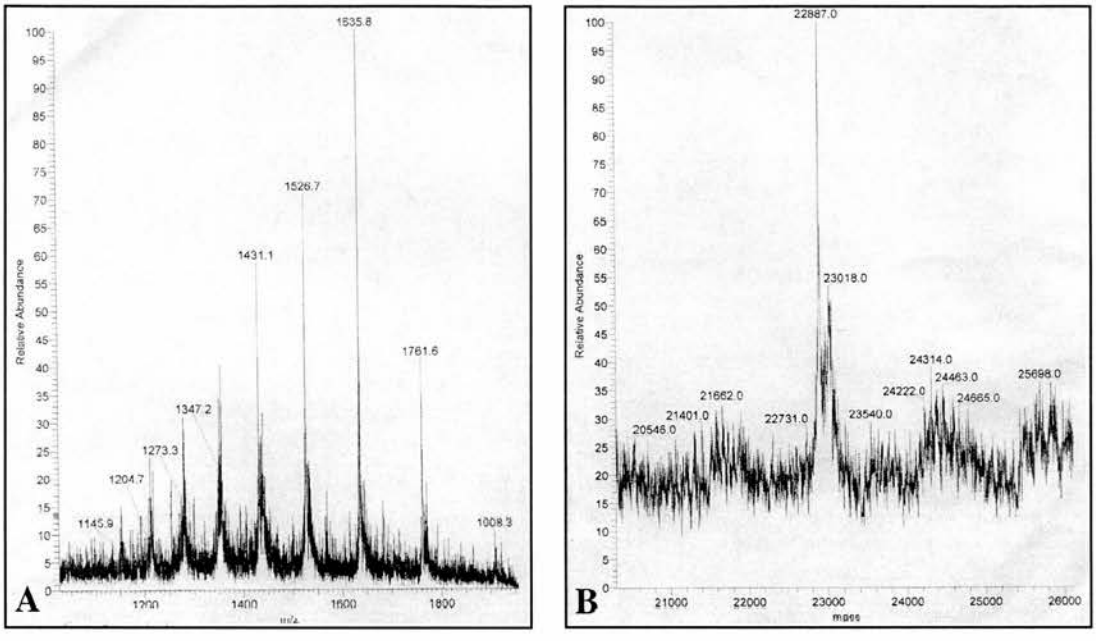


Figure 66

Results from mass spectrometric analysis of purified FKBP 22 from *N. crassa*.

Panel A, illustration of the peaks characterising the m/z ratios of the different multiply charged ions present in solution following electrospray ionisation mass spectrometry of FKBP 22. Panel B, automatic deconvolution of the m/z spectrum allowed identification of the molecular mass of proteins present in solution.

7.5.4

Crystallisation of FKBP 22

Two crystal forms were grown from the protein isolated using the purification protocol outlined in Section 7.2.3. The standard hanging drop vapour diffusion method, as described in Chapter 2, Section 2.5.1.3 was followed.

Photographs illustrating changes in crystal morphology are reproduced in Figure 67.

7.5.5

Preliminary characterisation and data collection

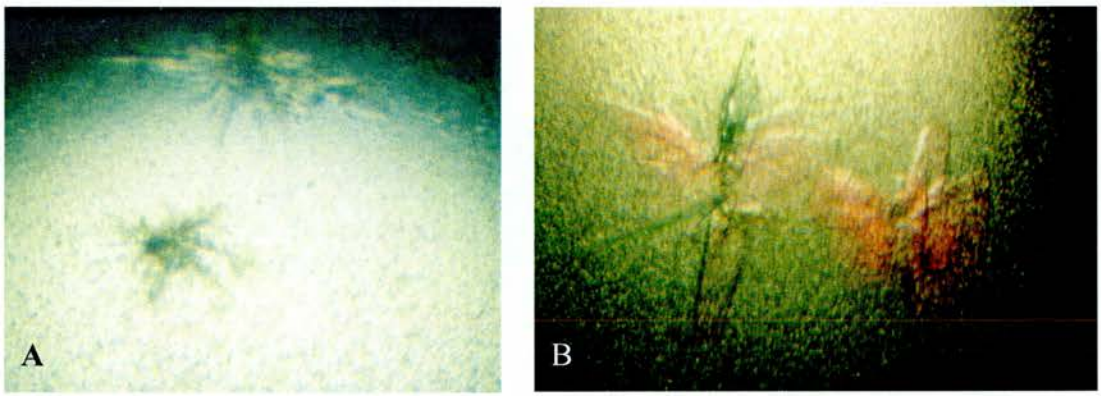


Figure 67 A and B

Photographs showing changes in FKBP 22 crystal morphology during the process of attempting refinement of crystallisation conditions

Clusters of Micro-needles (A) and Twinned thin plate crystals (B):

Crystals of the forms pictured above grew from the same mother liquor.

The well solution consisted of 600µl Methoxypolyethylene glycol 5000, 100 µl 1M 2-(N-Morpholino)ethane-sulphonic acid (MES) at pH 6.1, 30µl saturated ammonium sulphate, 10µl Dimethyl sulphoxide (DMSO) and 260 µl water. The drop consisted of 2µl protein solution and 2µl of well solution. The protein solution consisted of protein at either 6 or 12mg/ml in phosphate buffered saline at pH 7.0.

Crystals grew as very thin twinned/multiple sheets.

The protein crystallised very readily, producing numerous clusters of very small crystals. Data from a crystal such as that illustrated in Figure 67B was collected on ID 14, EM 3 at the ESRF, Grenoble. Despite a badly twinned diffraction pattern (Figure 68), it was possible to separate out a pattern corresponding to one single crystal. Data was processed using DENZO and Scalepack.

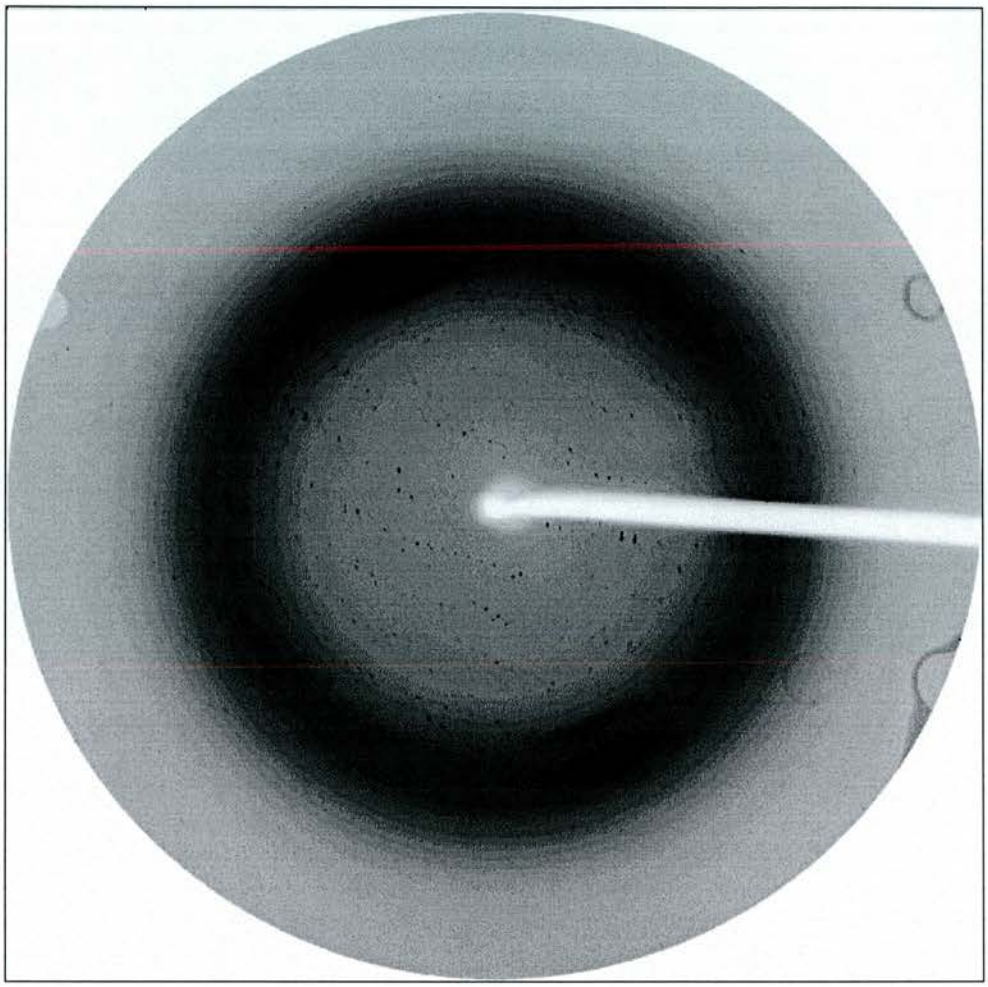


Figure 68

Typical diffraction pattern obtained from FKBP 22. The diffraction pattern shows data being produced from two crystals, an example of twinning.

Attempts at further refinement of these initial conditions, as outlined in the Materials and Methods Chapter, Section 2.5.1.3 had little effect on crystal morphology. Extensive refinement of crystallisation conditions has, as yet, failed to produce large single crystals capable of diffracting to high resolution.

7.6

FKBP 22 PARTIAL STRUCTURE

7.6.1

Structure solution and refinement

Crystals belong to space group $P2_1$, with unit cell dimensions a= 70.83 b= 32.32 c= 77.10 Å ,ß = 106.29^0. The merging R factor is 10.6% and the data is 86% complete to a resolution of 2.6Å.

A partial structure was determined, using the molecular replacement programme MOLREP (Vagin and Teplyakov, 1997), with a native FKBP 12 structure (Wilson et al., 1995) as the search model. There are two molecules within the asymmetric unit. The first two peaks of the rotation function produced peaks of Rf/sigma values of 6.63 and 5.93 (Figure 69). The translation function of the first molecule produced an R factor of 0.588 and a correlation factor of 0.185. The second molecule produced an R factor and correlation factor of 0.543 and 0.306 respectively with a sigma of 18.55.

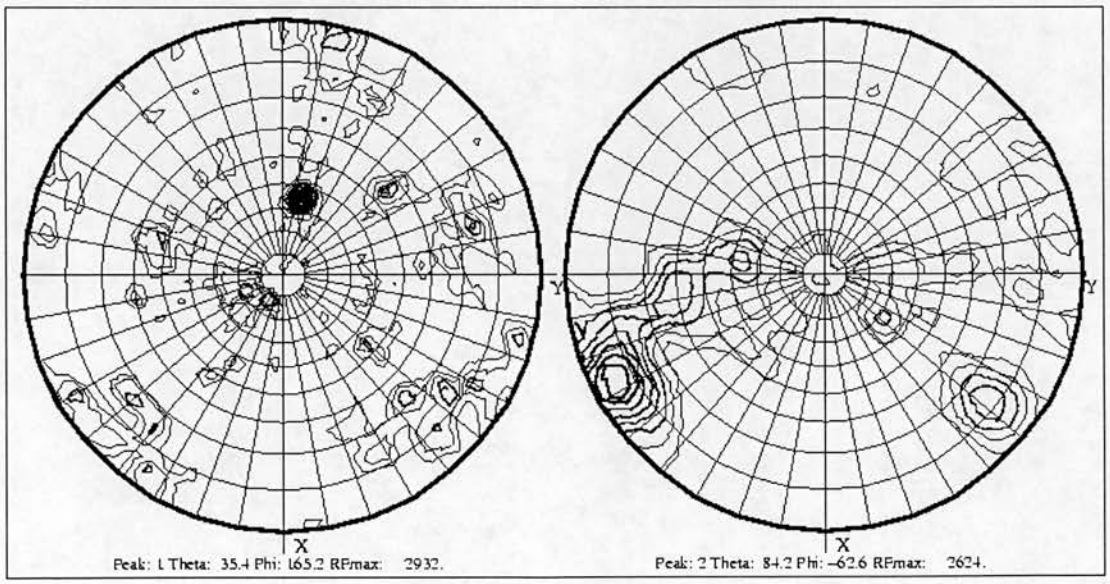


Figure 69

Representation of output from MOLREP rotation function. The first two peaks of the cross rotation function of the FKBP domains are shown.

The electron density for the FKBP domain of the structure is very clear, however the quality of data used did not allow unambiguous chain tracing extending into the C or N terminal regions.

7.6.1.1

Summary of crystallographic statistics for FKBP 22

Crystallographic statistics for FKBP 22	
Resolution range	30.0-2.6
Space group	$P2_1$
a	70.83
b	32.23
c	77.10
β	106.29
R_{merge}	0.106
(top shell)	0.281
Completeness %	87.0
(top shell)	82.0
No. of observations	20652
No. of reflections	9372

Table 19

Summary of crystallographic statistics used in the determination of the partial structure solution of FKBP 22

7.6.2

FKBP domain packing

The packing diagram for the FKBP domain (Figure 70) shows the FKBP domain packing efficiently. Analysis of close contacts from the crystal packing revealed very few bad contacts, only a few side chains were inappropriate. Such results are typical of those obtained from a molecular replacement solution. Packing of the FKBP

domain leaves large channels available for the remainder of the molecule, the approximately 10kDa C-terminal domain predicted to be helical.

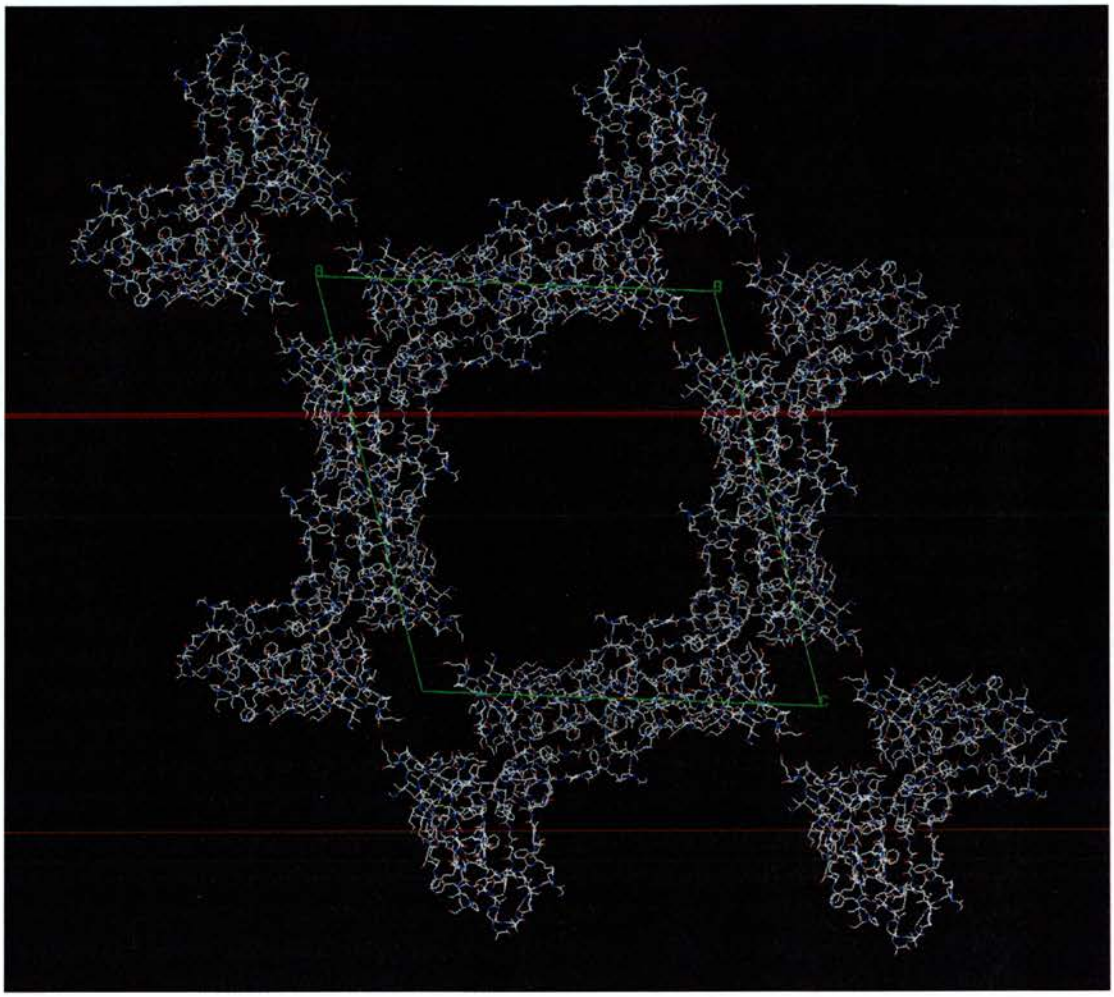


Figure 70

Packing diagram of the FKBP domain of FKBP 22. The unit cell is shown in green.

7.6.3

The overall architecture of FKBP 22

Secondary structure prediction suggested that residues 116-197 form an amphipathic α-helix (Solscheid and Tropschug, 2000). A similar prediction tool, PSIpred (http://insulin.brunel.ac.uk/psipred/), shows the C-terminal to be highly helical (Figure 71).

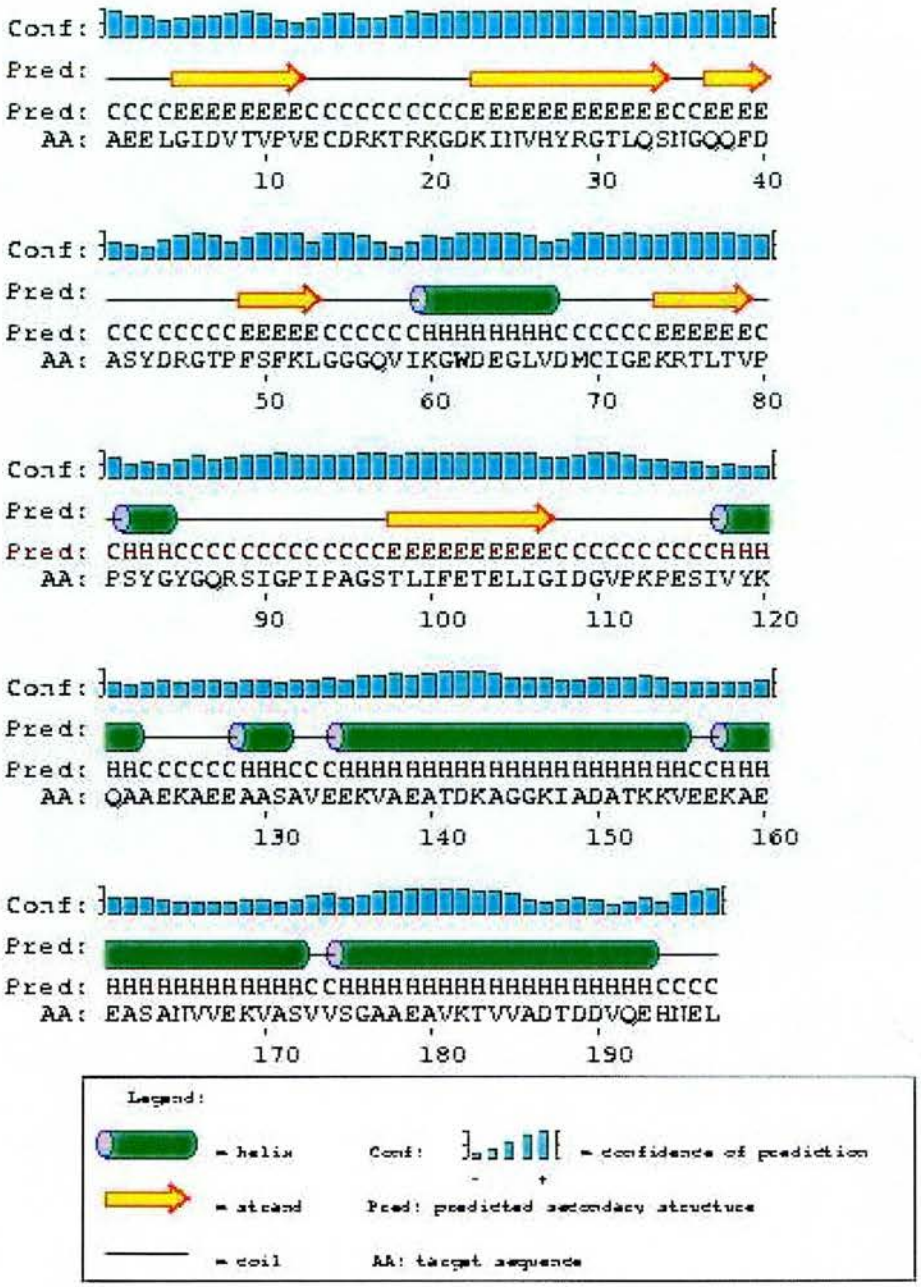


Figure 71

Output from the secondary structure prediction programme PSIpred. The C-terminal region of the protein is predicted to be highly helical.

Attempts to locate the predicted helices of the C-terminal domain using MOLREP (Vagin and Teplyakov, 1997) were unsuccessful. Manual chain tracing of the C-terminal domain was also tried, however data was not of sufficiently high quality to allow this technique to be useful.

Analysis of crystal packing shows there is clearly space within the molecule to accommodate the C-terminal domain, with possible helix-helix interactions between the two crystallographically related molecules.

The partial structure of the FKBP domain generated is very well conserved with that of all other FKBP proteins determined to date. This might have been expected due to the high degree of sequence identity shared between proteins over this domain. A BLAST search of the SWISS-PROT database revealed homologues of this protein to be present in *D. melanogaster* and *Homo sapiens*. Figure 72 illustrates a sequence alignment of these proteins with FKBP 22 from *N. crassa* and human FKBP 12.

```
FKBP12_HUMAN  ---------------------------------GVQVETIS-PGDGRTFPKRGQTCVVHY
FKBP23_DMELA  MSKSNLVISCLLLVAISNSLVRAQ---------DLKVEVISTPEVCEQKSKNGDSLTMHY
FKBP22_HUMAN  MPKTMHFLFRFIVFFYLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHY
FKBP22_NEUCR  --MKSIFLSLSLLASATVGVLAAE---------ELGIDVTV-PVECDRKTRKGDKINVHY
                                                : :::.   *      .:.*:    **

FKBP12_HUMAN  TGMLE-DGKKFDSSRDRNK--PFKFMLGKQEVIRGWEEGVAQMSVGQRAKLTISPDYAYG
FKBP23_DMELA  TGTLQADGKKFDSSFDRDQ--PFTFQLGAGQVIKGWDQGLLNMCVGEKRKLTIPPQLGYG
FKBP22_HUMAN  DGYLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDMCPGEKRKVVIPPSFAYG
FKBP22_NEUCR  RGTLQSNGQQFDASYDRGT--PFSFKLGGGQVIKGWDEGLVDMCIGEKRTLTVPPSYGYG
               * *  :*.:* .*  :.   *  * **  :**:* : .: :*. *:: .:.:.*. .**

FKBP12_HUMAN  ATGHP-GIIPPHATLVFDVELLKLE-----------------------------------
FKBP23_DMELA  DQGAG-NVIPPKATLLFDVELINIGNAP-PTTNVFKEIDDNADKQLS--REEVSEYLKKQ
FKBP22_HUMAN  KEGHAEGKIPPDATLIFEIELYAVTKGP-RSIETFKQIDMDNDRQLS--KAEINLYLQR-
FKBP22_NEUCR  QRSIG--PIPAGSTLIFETELIGIDGVPKPESIVYKQAAEKAEEAASAVEEKVAEATDKA
                .     **. :**:*: **  :

FKBP12_HUMAN  ---------------------------------------------------------
FKBP23_DMELA  MTAVEGQDSEELKNMLAENDK-LVEEIFQHED--KDKNGFISHDEFSGPKHDEL
FKBP22_HUMAN  ----EFEKDEKPRDK-SYQDA-VLEDIFKKND--HDGDGFISPKEYNVYQHDEL
FKBP22_NEUCR  G-GKIADATKKVEEKAEEASANVVEKVASVVSGAAEAVKTVVADTDDVQEHNEL
```

Figure 72

Alignment of the sequences of FKBP 22 from *N. crassa*, FKBP 23 from *D. melanogaster* and human FKBP 23 and FKBP 12, (SWISS-PROT accession numbers O60046, Q9V3V2, Q9Y6BO and P20071 respectively. The alignment was carried out using CLUSTALW.

Conservation of FKBP 22 between different species suggests an important role for this particular member of the FKBP family, closer sequence analysis studies coupled with improved structural data may shed more light onto the possible functions of this family member.

CHAPTER 8

SUMMARY AND FUTURE WORK

Results presented in this thesis include characterisation and structural studies on three members of the immunophilin family, Cyclophilin 3 from the free living nematode *C. elegans*, bovine Cyclophilin 40 and FKBP 22 from *N. crassa*.
This chapter will include a brief summary of results presented earlier, placing new data into context and will include some discussion of possible future directions for each of the projects.

CYP-3

The *C. elegans* genome sequencing project isolated 18 cyclophilins and 8 FKBPs, a significant proportion of this relatively small genome encoding for members of the immunophilin family. Such ubiquity suggests essential functions for these proteins, as yet largely to be determined. Cyclophilin 3 is one of the most abundantly expressed isoforms in *C. elegans*, and has been classified as a Cyclophilin A type protein. *C. elegans* is a well characterised organism, amenable to genetic analysis. The post-embryonic temporal expression pattern of *cyp-3* indicates that this transcript is most abundant during early larval development, peaking at the second larval stage and dropping off in later development. A very specific tissue localisation for this highly conserved cyclophilin was also observed, the protein expressed in the single excretory cell of *C. elegans*.

Possible roles of CYP-3 in the nematode were also discussed, temporal and spatial expression patterns for CYP-3 were characterised.

The high resolution X-ray structure of CYP-3 allowed comparisons with other known cyclophilin structures, resulting in identification of a "divergent loop" subclass. This subclass is defined by the following feature; an extra loop formed from inserted residues. Also frequently present in cyclophilins containing this "divergent loop", are two cysteines in close proximity (present in the structure in the

reduced form). Conserved histidine and glutamate residues may also be important, the glutamate acting to pin down the "divergent loop". A possible role in redox signal transduction or metal co-ordination was suggested for this protein. Formation of a disulphide bond between the two cysteines may possibly be inducing conformation change within the protein, resulting in a redox sensitive signalling mechanism. Results from experiments attempting to induce disulphide bond formation by metal catalysed oxidation were inconclusive, however crystals soaked with metals diffracted either very poorly or not at all, suggestive of metals perturbing the crystal lattice.

CYP-3 proved to be a very robust and useful system for ligand binding studies. Soaking crystals with the dipeptide Ala-Pro, coupled with solution studies, allowed correlation of a crystallographic binding constant with that determined from solution, the first report of such a study. Binding of a gold complex to the active site histidine of the protein was also reported.

Further experiments might include repeated attempts to induce oxidation of the two cysteine residues, to establish if the reaction does indeed induce localised conformation changes within the molecule or otherwise act in redox sensing signal transduction. The protein also provides an excellent target for ligand binding studies, due to the availability of crystals amenable to soaking, freezing and short data collection times. A number of other ligands could be investigated to further explore the specificity of binding.

CYCLOPHILIN 40

The high resolution X-ray structure of bovine Cyclophilin 40, the first reported structure of a two domain immunophilin; was discussed in Chapter 4. The structure revealed Cyp 40 to have a modular domain architecture in which the two domains were clearly separated into an N-terminal cyclophilin domain from residues 1 to 183, a linker region of 30 amino acids from residues 184 to 213 and a C-terminal TPR domain of 157 amino acids from residues 214 to 370. The Cyp domain can be classified as a member of the "divergent loop" subfamily, sharing the required

features of inserted residues forming a loop pinned down by a highly conserved glutamate residue and two cysteines present in the unoxidised form. The TPR domain consists largely of seven helices forming a right-handed super-helical structure, with the final C-terminal helix (the calmodulin binding helix), protruding beyond the shorter TPR helices.

A second crystal form produced a very different conformation that retained only the first TPR motif; the second TPR motif having sprung out to form an elongated helix of approximately 42Å. This crystal form potentially represents a trapped intermediate in the folding pathway of the fully folded monoclinic crystal form. Inter and intra molecular interactions are conserved between helices in the two structures, suggesting the possibility that tetragonal Cyp 40 may be an example of a domain swapped dimer.

Two proteins were isolated by bioaffinity interaction studies with Cyp 40; GAPDH and Hsp90. Further studies may identify additional interacting proteins.

The TPR domain in Cyp 40 was compared to other structurally characterised TPR proteins and was revealed to be structurally the most divergent from the "classical" helix turn helix motif first seen in the structure of PP5. Further analysis of protein sequences and structures allowed identification of residues specifically conserved between members of the Hsp90 binding subfamily of proteins, roles for individual residues in either maintaining the structural integrity of the molecule or in the interaction with Hsp90 were suggested.

Future work may include characterisation of Cyp 40 interactions with partner proteins including calmodulin and Hsp90.

The high resolution X-ray structure of bovine Cyp 40 was used as a template in automated model building of structures of five Cyp 40 homologues from diverse species. Resultant structures were compared with that of the experimentally derived bovine Cyp 40. The cyclophilin domains were very similar, mainly differing in

surface loop regions. Despite relatively low sequence identity, a remarkable degree of conservation of the overall TPR domain fold between Cyp 40 type proteins was observed. Areas of structural divergence between template and modelled structure were largely confined to N and C-terminal regions of individual TPR motifs. This suggested variable regions provided specificity of interaction for partner proteins, whilst conserved regions maintained the overall superhelical TPR domain structure, providing a binding interface for partner proteins.

Insights provided by Cyp 40 crystal growth and packing led to development of a modelled interaction for the five C-terminal residues of Hsp90 (MEEVD) within the Cyp 40 binding groove. This preliminary model was extended to include an additional 90 residues from the protein sequence. The model developed was of a three helix bundle interacting through the Cyp 40 binding groove. The potential range of interaction sites between TPR proteins and their physiological partner proteins is yet to be fully characterised, the modelled structure described above illustrates one of what may prove to be a number of potential sites of interaction with partner proteins.

The modelled structure could be further tested by site point mutagenesis studies of residues in both protein partners, followed by analysis of binding interactions between the two proteins. Such molecular biology experiments coupled with structural studies may prove a very powerful method by which to address defining the interaction between immunophilins and Hsp90. Studies including structural characterisation of a complex between Cyp 40 and the Hsp90 C-terminal domain fragments discussed in Chapter 5 are planned.

FKBP 22

The partial X-ray structure of FKBP 22 from *Neurospora crassa* was reported. This protein contains an N-terminal FKBP domain with a novel C-terminal domain. To date, only the FKBP domain has been characterised. Experiments are underway to try and improve crystal quality with the overall goal of solving the structure of the whole protein.

REFERENCE LIST

[Anon] (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science *282*, 2012-2018.

Abe,Y., Shodai,T., Muto,T., Mihara,K., Torii,H., Nishikawa,S., Endo,T., and Kohda,D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. Cell *100*, 551-560.

Adams,M.D., Celniker,S.E et al and Venter,J.C. (2000). The genome sequence of *Drosophila melanogaster*. Science *287*, 2185-2195.

Aligue,R., Akhavanniak,H., and Russel,P. (1994). A role for Hsp90 in cell-cycle control - Wee1 tyrosine kinase activity requires interaction with Hsp90. EMBO Journal *13*, 6099-6106.

Anderson,S.K., Gallinger,S., Roder,J., Frey,J., Young,H.A., and Ortaldo,J.R. (1993). A cyclophilin-related protein involved in the function of natural killer cells. Proceedings of the National Academy of Sciences of the United States of America *90*, 542-546.

Andreeva,L., Heads,R., and Green,C.J. (1999). Cyclophilins and their possible role in the stress response. International Journal of Experimental Pathology *80*, 305-315.

Argon,Y. and Simen,B.B. (1999). GRP94, an ER chaperone with protein and peptide binding properties. Seminars in Cell & Developmental Biology *10*, 495-505.

Aslund,F. and Beckwith,J. (1999). Bridge over troubled waters: Sensing stress by disulfide bond formation. Cell *96*, 751-753.

Bachinger,H.P. (1987). The influence of peptidyl-prolyl *cis-trans* isomerase on the *in vitro* folding of type-III collagen. Journal Of Biological Chemistry *262*, 17144-17148.

Bajorath,J., Stenkamp,R., and Aruffo,A. (1993). Knowledge-based model-building of proteins - Concepts and examples. Protein Science *2*, 1798-1810.

Baker,E.K., Colley,N.J., and Zuker,C.S. (1994). The cyclophilin homolog NinaA functions as a chaperone, forming a stable complex *in vivo* with its protein target rhodopsin. EMBO Journal *13*, 4886-4895.

Ban,C., Junop,M., and Yang,W. (1999). Transformation of MutL by ATP binding and hydrolysis: A switch in DNA mismatch repair. Cell *97*, 85-97.

Bang,H. and Fischer,G. (1991). Slow conformational changes in protein folding can be accelerated by enzymes. Biomedica Biochimica Acta *50*, S 137-S 142.

Bateman,A., Birney,E., Durbin,R., Eddy,S.R., Howe,K.L., and Sonnhammer,E.L.L. (2000). The Pfam protein families database. Nucleic Acids Research *28*, 263-266.

Baughman,G., Wiederrecht,G.J., Campbell,N.F., Martin,M.M., and Bourgeois,S. (1995). FKBP 51, a novel T cell-specific immunophilin capable of calcineurin inhibition. Molecular And Cellular Biology *15*, 4395-4402.

Bennett,M.J., Schlunegger,M.P., and Eisenberg,D. (1995). 3D domain swapping - A mechanism for oligomer assembly. Protein Science *4*, 2455-2468.

Berardini,T.Z., Bollman,K., Sun,H., and Poethig,R.S. (2001). Regulation of vegetative phase change in *Arabidopsis thaliana* by cyclophilin 40. Science *291*, 2405-2407.

Bierer,B.E., Mattila,P.S., Standaert,R.F., Herzenberg,L.A., Burakoff,S.J., Crabtree,G., and Schreiber,S.L. (1990). 2 distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. Proceedings of the National Academy of Sciences of the United States of America *87*, 9231-9235.

Blagosklonny,M.V., Toretsky,J., Bohen,S., and Neckers,L. (1996). Mutant conformation of p53 translated *in vitro* or *in vivo* requires functional HSP90. Proceedings of the National Academy of Sciences of the United States of America *93*, 8379-8383.

Blatch,G.L. and Lassle,M. (1999). The tetratricopeptide repeat: A structural motif mediating protein- protein interactions. Bioessays *21*, 932-939.

Borel,J.F., Dipadova,F., Mason,J., Quesniaux,V., Ryffel,B., and Wenger,R. (1989). Pharmacology of cyclosporine (sandimmune) .1. Introduction. Pharmacological Reviews *41*, 239-242.

Borkovich,K.A., Farrelly,F.W., Finkelstein,D.B., Taulien,J., and Lindquist,S. (1989). Hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. Molecular And Cellular Biology *9*, 3919-3930.

Bowie,J.U. (1997). Helix packing angle preferences. Nature Structural Biology *4*, 915-917.

Braaten,D. and Luban,J. (2001). Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells. EMBO Journal *20*, 1300-1309.

Bradford,M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry *72*, 248-254.

Braun,W., Kallen,J., Mikol,V., Walkinshaw,M.D., and Wuthrich,K. (1995). 3-dimensional structure and actions of immunosuppressants and their immunophilins. Faseb Journal *9*, 63-72.

Buchner,J. (1999). Hsp90 & Co. - A holding for folding. Trends In Biochemical Sciences *24*, 136-141.

Bultynck,G., De Smet,P., Rossi,D., Callewaert,G., Missiaen,L., Sorrentino,V., De Smedt,H., and Parys,J.B. (2001). Characterization and mapping of the 12 kDa FK506 binding protein (FKBP12) binding site on different isoforms of the ryanodine receptor and of the inositol 1,4,5-trisphosphate receptor. Biochemical Journal *354*, 413-422.

Burkhard,P., Taylor,P., and Walkinshaw,M.D. (2000). X-ray structures of small ligand FKBP complexes provide an estimate for hydrophobic interaction energies. Journal Of Molecular Biology *295*, 953-962.

Callebaut,I., Renoir,J.M., Lebeau,M.C., Massol,N., Burny,A., Baulieu,E.E., and Mornon,J.P. (1992). An immunophilin that binds m(r) 90,000 heat-shock protein - main structural features of a mammalian p59 protein. Proceedings of the National Academy of Sciences of the United States of America *89*, 6270-6274.

Cameron,A.M., Steiner,J.P., Sabatini,D.M., Kaplin,A.I., Walensky,L.D., and Snyder,S.H. (1995). Immunophilin FK506 binding protein associated with inositol 1,4, 5- trisphosphate receptor modulates calcium flux. Proceedings of the National Academy of Sciences of the United States Of America *92*, 1784-1788.

Carrello,A., Ingley,E., Minchin,R.F., Tsai,S., and Ratajczak,T. (1999). The common tetratricopeptide repeat acceptor site for steroid receptor associated immunophilins and Hop is located in the dimerization domain of Hsp90. Journal Of Biological Chemistry *274*, 2682-2689.

Chambraud,B., Radanyi,C., Camonis,J.H., Rajkowski,K., Schumacher,M., and Baulieu,E.E. (1999). Immunophilins, Refsum disease, and lupus nephritis: The peroxisomal enzyme phytanoyl-Coa alpha-hydroxylase is a new FKBP associated protein. Proceedings of the National Academy of Sciences of the United States of America *96*, 2104-2109.

Chappell,L.H. and Wastling,J.M. (1992). Cyclosporin A - antiparasite drug, modulator of the host- parasite relationship and immunosuppressant. Parasitology *105*, S 25-S 40.

Chen,C.F., Chen,Y.M., Dai,K., Chen,P.L., Riley,D.J., and Lee,W.H. (1996a). A new member of the Hsp90 family of molecular chaperones interacts with the retinoblastoma protein during mitosis and after heat shock. Molecular And Cellular Biology *16*, 4691-4699.

Chen,M.S., Silverstein,A.M., Pratt,W.B., and Chinkers,M. (1996b). The tetratricopeptide repeat domain of protein phosphatase 5 mediates binding to glucocorticoid receptor heterocomplexes and acts as a dominant negative mutant. Journal Of Biological Chemistry *271*, 32315-32320.

Chen,S.Y. and Smith,D.F. (1998). Hop as an adaptor in the heat shock protein 70 (Hsp70) and Hsp90 chaperone machinery. Journal Of Biological Chemistry *273*, 35194-35200.

Chen,S.Y., Sullivan,W.P., Toft,D.O., and Smith,D.F. (1998). Differential interactions of p23 and the TPR-containing proteins Hop, Cyp40, FKBP52 and FKBP51 with Hsp90 mutants. Cell Stress & Chaperones *3*, 118-129.

Cho,H., Ramaswamy,S., and Plapp,B.V. (1997). Flexibility of liver alcohol dehydrogenase in stereoselective binding of 3-butylthiolane 1-oxides. Biochemistry *36*, 382-389.

Chothia,C. and Lesk,A.M. (1986). The relation between the divergence of sequence and structure in proteins. EMBO Journal *5*, 823-826.

Chou,I.T. and Gasser,C.S. (1997). Characterization of the cyclophilin gene family of *Arabidopsis thaliana* and phylogenetic analysis of known cyclophilin proteins. Plant Molecular Biology *35*, 873-892.

Csermely,P., Schnaider,T., Soti,C., Prohaszka,Z., and Nardai,G. (1998). The 90kDa molecular chaperone family: Structure, function, and clinical applications. A comprehensive review. Pharmacology & Therapeutics *79*, 129-168.

Cutforth,T. and Rubin,G.M. (1994). Mutations in Hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in *Drosophila*. Cell *77*, 1027-1036.

Daothi,M.H., Transue,T.R., Pelle,R., Murphy,N.B., Poortmans,F., and Steyaert,J. (1998). Expression, purification, crystallization and preliminary X-ray analysis of cyclophilin A from the bovine parasite *Trypanosoma brucei brucei*. Acta Crystallographica Section D-Biological Crystallography *54*, 1046-1048.

Das,A.K., Cohen,P.T.W., and Barford,D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein protein interactions. EMBO Journal *17*, 1192-1199.

Dolinski,K., Muir,S., Cardenas,M., and Heitman,J. (1997). All cyclophilins and FK506 binding proteins are, individually and collectively, dispensable for viability in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America *94*, 13093-13098.

Dornan,J., Page,A.P., Taylor,P., Wu,S.Y., Winter,A.D., Husi,H., and Walkinshaw,M.D. (1999a). Biochemical and structural characterization of a divergent loop cyclophilin from *Caenorhabditis elegans*. Journal Of Biological Chemistry *274*, 34877-34883.

Dornan,J., Taylor,P., Carrello,A., Minchin,R.F., Ratajczak,T., and Walkinshaw,M.D. (1999b). Purification, characterisation and crystallisation in two crystal forms of bovine cyclophilin 40. Acta Cryst. D 1079-1082.

Doyle,V., Virji,S., and Crompton,M. (1999). Evidence that cyclophilin-A protects cells against oxidative stress. Biochemical Journal *341*, 127-132.

Dreher,D., Vargas,J.R., Hochstrasser,D.F., and Junod,A.F. (1995). Effects of oxidative stress and Ca2+ agonists on molecular chaperones in human umbilical vein endothelial cells. Electrophoresis *16*, 1205-1214.

Ducruix, A and Giege, R. Crystallization of Nucleic Acids and Proteins, A Practical Approach, Second edition, 1998.

Duina,A.A., Chang,H.C.J., Marsh,J.A., Lindquist,S., and Gaber,R.F. (1996b). A cyclophilin function in Hsp90 dependent signal transduction. Science *274*, 1713-1715.

Duina,A.A., Kalton,H.M., and Gaber,R.F. (1998a). Requirement for Hsp90 and a Cyp 40 type cyclophilin in negative regulation of the heat shock response. Journal Of Biological Chemistry *273*, 18974-18978.

Duina,A.A., Marsh,J.A., and Gaber,R.F. (1996a). Identification of 2 Cyp 40 like cyclophilins in *Saccharomyces cerevisiae*, one of which is required for normal growth. Yeast *12*, 943-952.

Duina,A.A., Marsh,J.A., Kurtz,R.B., Chang,H.C.J., Lindquist,S., and Gaber,R.F. (1998b). The peptidyl-prolyl isomerase domain of the Cyp 40 cyclophilin homolog Cpr7 is not required to support growth or glucocorticoid receptor activity in *Saccharomyces cerevisiae*. Journal Of Biological Chemistry *273*, 10819-10822.

Eberhardt,E.S., Loh,S.N., Hinck,A.P., and Raines,R.T. (1992). Solvent effects on the energetics of prolyl peptide bond isomerization. Journal Of The American Chemical Society *114*, 5437-5439.

Ellis,J. (1987). Proteins as Molecular Chaperones. Nature *328*, 378-379.

Ellis,R.J. and Hemmingsen,S.M. (1989). Molecular Chaperones - Proteins essential for the biogenesis of some macromolecular structures. Trends In Biochemical Sciences *14*, 339-342.

Esnouf,R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. Journal of Molecular Graphics & Modelling *15*, 132-&.

Faure,J.D., Vittorioso,P., Santoni,V., Fraisier,V., Prinsen,E., Barlier,I., Van Onckelen,H., Caboche,M., and Bellini,C. (1998). The PASTICCINO genes of *Arabidopsis thaliana* are involved in the control of cell division and differentiation. Development *125*, 909-918.

Felts,S.J., Owen,B.A.L., Nguyen,P., Trepel,J., Donner,D.B., and Toft,D.O. (2000). The Hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties. Journal Of Biological Chemistry *275*, 3305-3312.

Ferreira,P.A., Hom,J.T., and Pak,W.L. (1995). Retina specifically expressed novel subtypes of bovine cyclophilin. Journal Of Biological Chemistry *270*, 23179-23188.

Ferreira,P.A., Nakayama,T.A., Pak,W.L., and Travis,G.H. (1996). Cyclophilin related protein RanBp2 acts as chaperone for red/green opsin. Nature *383*, 637-640.

Ferreira,P.A., Nakayama,T.A., and Travis,G.H. (1997). Interconversion of red opsin isoforms by the cyclophilin-related chaperone protein Ran-binding protein 2. Proceedings of the National Academy of Sciences of the United States of America *94*, 1556-1561.

Fischer,G., Bang,H., and Mech,C. (1984). Detection of enzyme catalysis for *cis-trans* isomerization of peptide bonds using proline containing peptides as substrates. Biomedica Biochimica Acta *43*, 1101-1111.

Fischer,G., Berger,E., and Bang,H. (1989a). Kinetic beta deuterium isotope effects suggest a covalent mechanism for the protein folding enzyme peptidylprolyl *cis trans* isomerase. Febs Letters *250*, 267-270.

Fischer,G., Wittmannliebold,B., Lang,K., Kiefhaber,T., and Schmid,F.X. (1989b). Cyclophilin and peptidyl-prolyl *cis-trans* isomerase are probably identical proteins. Nature *337*, 476-478.

Franke,E.K. and Luban,J. (1996). Inhibition of HIV-1 replication by Cyclosporin A or related compounds correlates with the ability to disrupt the Gag- Cyclophilin A interaction. Virology *222*, 279-282.

Franke,E.K., Yuan,H.E.H., and Luban,J. (1994). Specific incorporation of Cyclophilin A into HIV-1 virions. Nature *372*, 359-362.

Fransson,C., Freskgard,P.O., Herbertsson,H., Johansson,A., Jonasson,P., Martensson,L.G., Svensson,M., Jonsson,B.H., and Carlsson,U. (1992). *Cis-trans* isomerization is rate determining in the reactivation of denatured human carbonic anhydrase II as evidenced by proline isomerase. Febs Letters *296*, 90-94.

Freeman,B.C. and Morimoto,R.I. (1996). The human cytosolic molecular chaperones Hsp90, Hsp70 (Hsc70) and Hdj-1 have distinct roles in recognition of a nonnative protein and protein refolding. EMBO Journal *15*, 2969-2979.

Freeman,B.C., Toft,D.O., and Morimoto,R.I. (1996). Molecular chaperone machines - chaperone activities of the cyclophilin Cyp 40 and the steroid aporeceptor associated protein p23. Science *274*, 1718-1720.

Freskgard,P.O., Bergenhem,N., Jonsson,B.H., Svensson,M., and Carlsson,U. (1992). Isomerase and chaperone activity of prolyl isomerase in the folding of carbonic anhydrase. Science *258*, 466-468.

Frydman,J. and Hohfeld,J. (1997). Chaperones get in touch: The Hip-Hop connection. Trends In Biochemical Sciences *22*, 87-92.

Gaburjakova,M., Gaburjakova,J., Reiken,S., Huang,F., Marx,S.O., Rosemblit,N., and Marks,A.R. (2001). FKBP 12 binding modulates ryanodine receptor channel gating. Journal Of Biological Chemistry *276*, 16931-16935.

Galat,A. and Metcalfe,S.M. (1995). Peptidylproline *cis/trans* isomerases. Progress In Biophysics & Molecular Biology *63*, 67-118.

Galigniana,M.D., Radanyi,C., Renoir,J.M., Housley,P.R., and Pratt,W.B. (2001). Evidence that the peptidylprolyl isomerase domain of the Hsp90 binding immunophilin FKBP 52 is involved in both dynein interaction and glucocorticoid receptor movement to the nucleus. Journal Of Biological Chemistry *276*, 14884-14889.

Gamble,T.R., Vajdos,F.F., Yoo,S.H., Worthylake,D.K., Houseweart,M., Sundquist,W.I., and Hill,C.P. (1996). Crystal structure of human cyclophilin A bound to the amino terminal domain of HIV-1 capsid. Cell *87*, 1285-1294.

Gatto,G.J., Geisbrecht,B.V., Gould,S.J., and Berg,J.M. (2000). Peroxisomal targeting signal 1 recognition by the TPR domains of human PEX5. Nature Structural Biology *7*, 1091-1095.

Ghirlanda,G., Lear,J.D., Lombardi,A., and DeGrado,W.F. (1998). From synthetic coiled coils to functional proteins: Automated design of a receptor for the calmodulin binding domain of calcineurin. Journal Of Molecular Biology *281*, 379-391.

Gidalevitz,T., Simen,B.B., Vogen,S.M., Dul,J.L., Case,C.M., and Argon,Y. (2001). GRP94 activity is essential for efficient protein folding in the ER. Biophysical Journal *80*, 1758.

Gothel,S.F. and Marahiel,M.A. (1999). Peptidyl-prolyl *cis-trans* isomerases, a superfamily of ubiquitous folding catalysts. Cellular and Molecular Life Sciences *55*, 423-436.

Gothel,S.F., Scholz,C., Schmid,F.X., and Marahiel,M.A. (1998). Cyclophilin and trigger factor from *Bacillus subtilis* catalyze *in vitro* protein folding and are necessary for viability under starvation conditions. Biochemistry *37*, 13392-13399.

Graumann,J., Lilie,H., Tang,X.L., Tucker,K.A., Hoffmann,J.T., Vijayalakshmi,J., Saper,M., Bardwell,J.C.A., and Jakob,U. (2001). Activation of the redox regulated molecular chaperone Hsp33 - A two step mechanism. Structure *9*, 377-387.

Grenert,J.P., Sullivan,W.P., Fadden,P., Haystead,T.A.J., Clark,J., Mimnaugh,E., Krutzsch,H., Ochel,H.J., Schulte,T.W., Sausville,E., Neckers,L.M., and Toft,D.O. (1997). The amino terminal domain of heat shock protein 90 (Hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates Hsp90 conformation. Journal Of Biological Chemistry *272*, 23843-23850.

Grizot,S., Fieschi,F., Dagher,M.C., and Pebay-Peyroula,E. (2001). The active N-terminal region of p67(phox) - Structure at 1.8 angstrom resolution and biochemical characterizations of the A128V mutant implicated in chronic granulomatous disease. Journal Of Biological Chemistry *276*, 21627-21631.

Groves,M.R., Hanlon,N., Turowski,P., Hemmings,B.A., and Barford,D. (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. Cell *96*, 99-110.

Guex,N., Diemand,A., and Peitsch,M.C. (1999). Protein modelling for all. Trends In Biochemical Sciences *24*, 364-367.

Guex,N. and Peitsch,M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis *18*, 2714-2723.

Gupta,R.S. (1995). Phylogenetic analysis of the 90 kDa heat-shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. Molecular Biology and Evolution *12*, 1063-1073.

Hamilton,G.S. and Steiner,J.P. (1998). Immunophilins: Beyond immunosuppression. Journal Of Medicinal Chemistry *41*, 5119-5143.

Handschumacher,R.E., Harding,M.W., Rice,J., and Drugge,R.J. (1984). Cyclophilin A specific cytosolic binding-protein for Cyclosporin A. Science *226*, 544-547.

Harding,M.W., Galat,A., Uehling,D.E., and Schreiber,S.L. (1989). A receptor for the immunosuppressant FK506 is a *cis-trans* peptidyl prolyl isomerase. Nature *341*, 758-760.

Harding,M.W., Handschumacher,R.E., and Speicher,D.W. (1986). Isolation and amino acid sequence of Cyclophilin. Journal Of Biological Chemistry *261*, 8547-8555.

Harrison,R.K. and Stein,R.L. (1990). Mechanistic studies of peptidyl prolyl *cis trans* isomerase - evidence for catalysis by distortion. Biochemistry *29*, 1684-1689.

Hartson,S.D., Irwin,A.D., Shao,J.Y., Scroggins,B.T., Volk,L., Huang,W.J., and Matts,R.L. (2000). p50(cdc37) is a nonexclusive Hsp90 cohort which participates intimately in Hsp90 mediated folding of immature kinase molecules. Biochemistry *39*, 7631-7644.

Helekar,S.A. and Patrick,J. (1997). Peptidyl prolyl cis-trans isomerase activity of cyclophilin A in functional homo-oligomeric receptor expression. Proceedings of the National Academy of Sciences of United States of America *94*, 5432-5437.

Helmbrecht,K., Zeise,E., and Rensing,L. (2000). Chaperones in cell cycle regulation and mitogenic signal transduction: A review. Cell Proliferation *33*, 341-365.

Hickey,E., Brandon,S.E., Smale,G., Lloyd,D., and Weber,L.A. (1989). Sequence and regulation of a gene encoding a human 89 kilodalton heat shock protein. Molecular And Cellular Biology *9*, 2615-2626.

Higgins,D.G., Thompson,J.D., and Gibson,T.J. (1996). Using CLUSTAL for multiple sequence alignments. Methods In Enzymology *266*, 383-402.

Hirano,T., Kinoshita,N., Morikawa,K., and Yanagida,M. (1990). Snap helix with knob and hole - essential repeats in *S. pombe* nuclear protein nuc2+. Cell *60*, 319-328.

Hirtzlin,J., Farber,P.M., Franklin,R.M., and Bell,A. (1995). Molecular and biochemical characterization of a *Plasmodium falciparum* cyclophilin containing a cleavable signal sequence. European Journal Of Biochemistry *232*, 765-772.

Hoerauf,A., Rascher,C., Bang,R., Pahl,A., Solbach,W., Brune,K., Rollinghoff,M., and Bang,H. (1997). Host cell cyclophilin is important for the intracellular replication of *Leishmania major*. Molecular Microbiology *24*, 421-429.

Hueros,G., Rahfeld,J., Salamini,F., and Thompson,R. (1998). A maize FK506-sensitive immunophilin, mzFKBP 66, is a peptidylproline *cis-trans* isomerase that interacts with calmodulin and a 36kDa cytoplasmic protein. Planta *205*, 121-131.

Ivery,M.T.G. (2000). Immunophilins: Switched on protein binding domains? Medicinal Research Reviews *20*, 452-484.

Jakob,U., Lilie,H., Meyer,I., and Buchner,J. (1995). Transient interaction of Hsp90 with early unfolding intermediates of citrate synthase - implications for heat shock *in-vivo*. Journal Of Biological Chemistry *270*, 7288-7294.

Jakob,U., Muse,W., Eser,M., and Bardwell,J.C.A. (1999). Chaperone activity with a redox switch. Cell *96*, 341-352.

Janowski,R., Kozak,M., Jankowska,E., Grzonka,Z., Grubb,A., Abrahamson,M., and Jaskolski,M. (2001). Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. Nature Structural Biology *8*, 316-320.

Jaschke,A., Mi,H.F., and Tropschug,M. (1998). Human T cell cyclophilin18 binds to thiol specific antioxidant protein Aop1 and stimulates its activity. Journal Of Molecular Biology *277*, 763-769.

Jayaraman,T., Brillantes,A.M., Timerman,A.P., Fleischer,S., Erdjumentbromage,H., Tempst,P., and Marks,A.R. (1992). FK506 binding protein associated with the calcium release channel (ryanodine receptor). Journal Of Biological Chemistry *267*, 9474-9477.

Jin,Z.G., Melaragno,M.G., Liao,D.F., Yan,C., Haendeler,J., Suh,Y.A., Lambeth,J.D., and Berk,B.C. (2000). Oxidative stress regulates Cyclophilin A function *in vitro* and *in vivo*. Circulation *102*, 154.

Johnson,B.D., Schumacher,R.J., Ross,E.D., and Toft,D.O. (1998). Hop modulates Hsp70/Hsp90 interactions in protein folding. Journal Of Biological Chemistry *273*, 3679-3686.

Johnson,J.L. and Craig,E.A. (1997). Protein folding *in vivo*: Unraveling complex pathways. Cell *90*, 201-204.

Jones,T.A., Zou,J.Y., Cowan,S.W., and Kjeldgaard,M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallographica Section A-Fundamentals Of Crystallography *47*, 110-119.

Kabsch,W. and Sander,C. (1983). Dictionary of protein secondary structure - pattern recognition of hydrogen bonded and geometrical features. Biopolymers *22*, 2577-2637.

Kallen,J., Spitzfaden,C., Zurini,M.G.M., Wider,G., Widmer,H., Wuthrich,K., and Walkinshaw,M.D. (1991). Structure of human Cyclophilin and its binding site for Cyclosporin A determined by X-ray crystallography and NMR spectroscopy. Nature *353*, 276-279.

Kampranis,S.C., Bates,A.D., and Maxwell,A. (1999). A model for the mechanism of strand passage by DNA gyrase. Proceedings Of The National Academy Of Sciences Of The United States Of America *96*, 8414-8419.

Kawai,J., Shinagawa,A and Hayashizaki,Y. (2001). Functional annotation of a full-length mouse cDNA collection. Nature *409*, 685-690.

Kay,J.E. (1996). Structure function relationships in the FK506 binding protein (FKBP) family of peptidylprolyl *cis-trans* isomerases. Biochemical Journal *314*, 361-385.

Ke,H.M. (1992). Similarities and differences between human Cyclophilin A and other beta-barrel structures - structural refinement at 1.63 angstrom resolution. Journal Of Molecular Biology *228*, 539-550.

Ke,H.M., Mayrose,D., and Cao,W. (1993). Crystal structure of Cyclophilin A complexed with substrate Ala- Pro suggests a solvent-assisted mechanism of *cis-trans* isomerization. Proceedings Of The National Academy Of Sciences Of The United States Of America *90*, 3324-3328.

Ke,H.M., Zydowsky,L.D., Liu,J., and Walsh,C.T. (1991). Crystal-structure of recombinant human T cell Cyclophilin A at 2.5-a resolution. Proceedings Of The National Academy Of Sciences Of The United States Of America *88*, 9483-9487.

Kern,D., Drakenberg,T., Wikstrom,M., Forsen,S., Bang,H., and Fischer,G. (1993). The *cis/trans* interconversion of the calcium regulating hormone calcitonin is catalyzed by Cyclophilin. Febs Letters *323*, 198-202.

Kern,G., Kern,D., Schmid,F.X., and Fischer,G. (1994). Reassessment of the putative chaperone function of prolyl- *cis/trans* isomerases. Febs Letters *348*, 145-148.

Kieffer,L.J., Thalhammer,T., and Handschumacher,R.E. (1992). Isolation and characterization of a 40kDa cyclophilin related protein. Journal Of Biological Chemistry *267*, 5503-5507.

Kim,S.J., Jeong,D.G., Chi,S.W., Lee,J.S., and Ryu,S.E. (2001). Crystal structure of proteolytic fragments of the redox sensitive Hsp33 with constitutive chaperone activity. Nature Structural Biology *8*, 459-466.

Kimmins,S. and MacRae,T.H. (2000). Maturation of steroid receptors: an example of functional cooperation among molecular chaperones and their associated proteins. Cell Stress & Chaperones *5*, 76-86.

Kino,T., Hatanaka,H., Hashimoto,M., Nishiyama,M., Goto,T., Okuhara,M., Kohsaka,M., Aoki,H., And Imanaka,H. (1987a). FK506, A novel immunosuppressant isolated from a *Streptomyces* .1. Fermentation, isolation, and physicochemical and biological characteristics. Journal Of Antibiotics *40*, 1249-1255.

Kino,T., Hatanaka,H., Miyata,S., Inamura,N., Nishiyama,M., Yajima,T., Goto,T., Okuhara,M., Kohsaka,M., Aoki,H., and Ochiai,T. (1987b). FK506, a novel immunosuppressant isolated from a *Streptomyces* .2. Immunosuppressive effect of FK506 *in vitro*. Journal Of Antibiotics *40*, 1256-1265.

Kofron,J.L., Kuzmic,P., Kishore,V., Colonbonilla,E., and Rich,D.H. (1991). Determination of kinetic constants for peptidyl prolyl *cis trans* isomerases by an improved spectrophotometric assay. Biochemistry *30*, 6127-6134.

Kraulis,P.J. (1991). Molscript - a program to produce both detailed and schematic plots of protein structures. Journal Of Applied Crystallography *24*, 946-950.

Kumar,A., Roach,C., Hirsh,I.S., Turley,S., deWalque,S., Michels,P.A.M., and Hol,W.G.J. (2001). An unexpected extended conformation for the third TPR motif of the peroxin PEX5 from *Trypanosoma brucei*. Journal Of Molecular Biology *307*, 271-282.

Kurek,I., Aviezer,K., Erel,N., Herman,E., and Breiman,A. (1999). The wheat peptidyl prolyl *cis-trans* isomerase FKBP 77 is heat induced and developmentally regulated. Plant Physiology *119*, 693-703.

Lang,K., Schmid,F.X., and Fischer,G. (1987b). Catalysis of protein folding by proline isomerase. Biological Chemistry Hoppe-Seyler *368*, 1070-1071.

Lang,K., Schmid,F.X., and Fischer,G. (1987a). Catalysis of protein folding by prolyl isomerase. Nature *329*, 268-270.

Lapouge,K., Smith,S.J.M., Walker,P.A., Gamblin,S.J., Smerdon,S.J., and Rittinger,K. (2000). Structure of the TPR domain of p67(phox) in complex with Rac center dot GTP. Molecular Cell *6*, 899-907.

Lebeau,M.C., Massol,N., Herrick,J., Faber,L.E., Renoir,J.M., Radanyi,C., and Baulieu,E.E. (1992). P59, an Hsp90 binding protein - cloning and sequencing of its cDNA and preparation of a peptide directed polyclonal antibody. Journal Of Biological Chemistry *267*, 4281-4284.

Leto,T.L. (1996). The p67-phox NADPH oxidase component is comprised of several tetratricopeptide repeats (TPR) that are disrupted in a rare form of chronic granulomatous disease. Faseb Journal *10*, 1516-1516.

Leverson,J.D. and Ness,S.A. (1998). Point mutations in v-myb disrupt a cyclophilin catalyzed negative regulatory mechanism. Molecular Cell *1*, 203-211.

Liao,D.F., Jin,Z.G., Baas,A.S., Daum,G., Gygi,S.P., Aebersold,R., and Berk,B.C. (2000). Purification and identification of secreted oxidative stress induced factors from vascular smooth muscle cells. Journal Of Biological Chemistry *275*, 189-196.

Lindquist,S. and Craig,E.A. (1988). The heat shock proteins. Annual Review of Genetics *22*, 631-677.

Liu,J., Albers,M.W., Chen,C.M., Schreiber,S.L., and Walsh,C.T. (1990). Cloning, expression, and purification of human cyclophilin in *Escherichia coli* and assessment of the catalytic role of cysteines by site directed mutagenesis. Proceedings Of The National Academy Of Sciences Of The United States Of America *87*, 2304-2308.

Liu,J., Chen,C.M., and Walsh,C.T. (1991b). Human and *Escherichia coli* cyclophilins - sensitivity to inhibition by the immunosuppressant Cyclosporin A correlates with a specific tryptophan residue. Biochemistry *30*, 2306-2310.

Liu,J., Farmer,J.D., Lane,W.S., Friedman,J., Weissman,I., and Schreiber,S.L. (1991a). Calcineurin is a common target of cyclophilin Cyclosporin A and FKBP-FK506 complexes. Cell *66*, 807-815.

Liu,J. and Walsh,C.T. (1990). Peptidyl prolyl cis-trans-isomerase from *Escherichia coli* - a periplasmic homolog of cyclophilin that is not inhibited by Cyclosporin A. Proceedings Of The National Academy Of Sciences Of The United States Of America *87*, 4028-4032.

Liu,Y.S., Gotte,G., Libonati,M., and Eisenberg,D. (2001). A domain-swapped RNase A dimer with implications for amyloid formation. Nature Structural Biology *8*, 211-214.

Lodish,H.F. and Kong,N. (1991). Cyclosporin A inhibits an initial step in folding of transferrin within the endoplasmic-reticulum. Journal Of Biological Chemistry *266*, 14835-14838.

Luban,J., Bossolt,K.L., Franke,E.K., Kalpana,G.V., and Goff,S.P. (1993). Human immunodeficiency virus type-1 Gag protein binds to Cyclophilin A and Cyclophilin-B. Cell *73*, 1067-1078.

Ma,Q. and Whitlock,J.P. (1997). A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Journal Of Biological Chemistry *272*, 8878-8884.

Macarthur,M.W. and Thornton,J.M. (1991). Influence of proline residues on protein conformation. Journal Of Molecular Biology *218*, 397-412.

Marcu,M.G., Chadli,A., Bouhouche,I., Catelli,M., and Neckers,L.M. (2000). The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the carboxyl terminus of the chaperone. Journal Of Biological Chemistry *275*, 37181-37186.

Mark,P.J., Ward,B.K., Kumar,P., Lahooti,H., Minchin,R.F., and Ratajczak,T. (2001). Human cyclophilin 40 is a heat shock protein that exhibits altered intracellular localization following heat shock. Cell Stress & Chaperones *6*, 59-70.

Maruya,M., Sameshima,M., Nemoto,T., and Yahara,I. (1999). Monomer arrangement in Hsp90 dimer as determined by decoration with N and C-terminal region specific antibodies. Journal Of Molecular Biology *285*, 903-907.

Massol,N., Lebeau,M.C., Renoir,J.M., Faber,L.E., and Baulieu,E.E. (1992). Rabbit FKBP59-heat shock protein binding immunophillin (HBI) is a calmodulin binding protein. Biochemical And Biophysical Research Communications *187*, 1330-1335.

Mayr,C., Richter,K., Lilie,H., and Buchner,J. (2000). Cpr6 and Cpr7, two closely related Hsp90 associated immunophilins from *Saccharomyces cerevisiae*, differ in their functional properties. Journal Of Biological Chemistry *275*, 34140-34146.

McGuffin,L.J., Bryson,K., and Jones,D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404-405.

McPherson, A. (1999). Crystallization of Biological Macromolecules. Cold Spring Harbour Laboratory Press.

McPherson,A., Malkin,A.J., and Kuznetsov,Y.G. (1995). The science of macromolecular crystallization. Structure *3*, 759-768.

Merritt,E.A. and Bacon,D.J. (1997). Raster3D: Photorealistic molecular graphics. Macromolecular Crystallography, Pt B *277*, 505-524.

MezaZepeda,L.A., Baudo,M.M., Palva,E.T., and Heino,P. (1998). Isolation and characterization of a cDNA corresponding to a stress activated cyclophilin gene in *Solanum commersonii*. Journal Of Experimental Botany *49*, 1451-1452.

Michnick,S.W., Rosen,M.K., Wandless,T.J., Karplus,M., and Schreiber,S.L. (1991). Solution structure of FKBP, a rotamase enzyme and receptor for FK506 and rapamycin. Science *252*, 836-839.

Mikol,V., Ma,D., and Carlow,C.K.S. (1998). Crystal structure of the cyclophilin like domain from the parasitic nematode *Brugia malayi*. Protein Science *7*, 1310-1316.

Moir,R.D., SethyCoraci,I., Puglia,K., Librizzi,M.D., and Willis,I.M. (1997). A tetratricopeptide repeat mutation in yeast transcription factor IIIC131 (TFIIIC131)

facilitates recruitment of TFIIB-related factor TFIIIB70. Molecular And Cellular Biology *17*, 7119-7125.

Morris,R.E. (1995). Mechanisms of action of new immunosuppressive drugs. Therapeutic Drug Monitoring *17*, 564-569.

Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K., and Ohara,O. (2000). Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. DNA Research *7*, 347-355.

Nair,S.C., Rimerman,R.A., Toran,E.J., Chen,S.Y., Prapapanich,V., Butts,R.N., and Smith,D.F. (1997). Molecular cloning of human FKBP51 and comparisons of immunophilin interactions with Hsp90 and progesterone receptor. Molecular And Cellular Biology *17*, 594-603.

Nathan,D.F. and Lindquist,S. (1995). Mutational analysis of Hsp90 function - interactions with a steroid receptor and a protein kinase. Molecular And Cellular Biology *15*, 3917-3925.

Nathan,D.F., Vos,M.H., and Lindquist,S. (1997). *In vivo* functions of the *Saccharomyces cerevisiae* Hsp90 chaperone. Proceedings Of The National Academy Of Sciences Of The United States Of America *94*, 12949-12956.

Navaza,J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr D. *A50*, 157-163.

Nelson,F.K. and Riddle,D.L. (1984). Functional study of the *Caenorhabditis elegans* secretory-excretory system using laser microsurgery. J. Exp. Zool. *231*, 45-46.

Nemoto,T., Oharanemoto,Y., Ota,M., Takagi,T., and Yokoyama,K. (1995). Mechanism of dimer formation of the 90-kDa heat-shock protein. European Journal Of Biochemistry *233*, 1-8.

Obermann,W.M.J., Sondermann,H., Russo,A.A., Pavletich,N.P., and Hartl,F.U. (1998). *In vivo* function of Hsp90 is dependent on ATP binding and ATP hydrolysis. Journal Of Cell Biology *143*, 901-910.

Ondek,B., Hardy,R.W., Baker,E.K., Stamnes,M.A., Shieh,B.H., and Zuker,C.S. (1992). Genetic dissection of cyclophilin function - saturation mutagenesis of the *Drosophila* cyclophilin homolog ninaA. Journal Of Biological Chemistry *267*, 16460-16466.

Otwinowski,Z. and Minor,W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods In Enzymology *276*, 307-326.

Ow,W.B., Luo,W., Park,Y.D., and Zhou,H.M. (2001). Chaperone like activity of peptidyl prolyl *cis-trans* isomerase during creatine kinase refolding. Protein Science *10*, 2346-2353.

Owensgrillo,J.K., Hoffmann,K., Hutchison,K.A., Yem,A.W., Deibel,M.R., Handschumacher,R.E., and Pratt,W.B. (1995). The Cyclosporin A binding immunophilin Cyp 40 and the FK506 binding immunophilin Hsp56 bind to a common site on Hsp90 and exist in independent cytosolic heterocomplexes with the untransformed glucocorticoid receptor. Journal Of Biological Chemistry *270*, 20479-20484.

Page,A.P., Hamilton,A.J., and Maizels,R.M. (1992b). *Toxocara canis* monoclonal antibodies to carbohydrate epitopes of secreted (tes) antigens localize to different secretion related structures in infective larvae. Experimental Parasitology *75*, 56-71.

Page,A.P., Kumar,S., and Carlow,C.K.S. (1995b). Parasite cyclophilins and anti-parasite activity of Cyclosporin A. Parasitology Today *11*, 385-388.

Page,A.P., Landry,D., Wilson,G.G., and Carlow,C.K.S. (1995a). Molecular characterization of a Cyclosporin A insensitive cyclophilin from the parasitic nematode *Brugia malayi*. Biochemistry *34*, 11545-11550.

Page,A.P., Macniven,K., and Hengartner,M.O. (1996). Cloning and biochemical characterization of the cyclophilin homologs from the free-living nematode *Caenorhabditis elegans*. Biochemical Journal *317*, 179-185.

Page,A.P., Rudin,W., Fluri,E., Blaxter,M.L., and Maizels,R.M. (1992a). *Toxocara canis* - a labile antigenic surface coat overlying the epicuticle of infective larvae. Experimental Parasitology *75*, 72-86.

Page,A.P. and Winter,A.D. (1998). A divergent multi domain cyclophilin is highly conserved between parasitic and free-living nematode species and is important in larval muscle development. Molecular And Biochemical Parasitology *95*, 215-227.

Panaretou,B., Prodromou,C., Roe,S.M., Obrien,R., Ladbury,J.E., Piper,P.W., and Pearl,L.H. (1998). ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone *in vivo*. EMBO Journal *17*, 4829-4836.

Papageorgiou,C., Sanglier,J.J., and Traber,R. (1996). Anti HIV-1 activity of a hydrophilic cyclosporin derivative with improved binding affinity to Cyclophilin A. Bioorganic & Medicinal Chemistry Letters *6*, 23-26.

Partaledis,J.A. and Berlin,V. (1993). The FKB2 gene of *Saccharomyces cerevisiae*, encoding the immunosuppressant binding protein FKBP 13, is regulated in response to accumulation of unfolded proteins in the endoplasmic reticulum. Proceedings Of The National Academy Of Sciences Of The United States Of America *90*, 5450-5454.

Pearl,L.H. and Prodromou,C. (2000). Structure and *in vivo* function of Hsp90. Current Opinion In Structural Biology *10*, 46-51.

Peattie,D.A., Harding,M.W., Fleming,M.A., Decenzo,M.T., Lippke,J.A., Livingston,D.J., and Benasutti,M. (1992). Expression and characterization of human FKBP 52, an immunophilin that associates with the 90kDa heat-shock protein and is

a component of steroid receptor complexes. Proceedings Of The National Academy Of Sciences Of The United States Of America *89*, 10974-10978.

Pflugl,G., Kallen,J., Schirmer,T., Jansonius,J.N., Zurini,M.G.M., and Walkinshaw,M.D. (1993). X-ray structure of a decameric cyclophilin cyclosporine crystal complex. Nature *361*, 91-94.

Pflugl,G.M., Kallen,J., Jansonius,J.N., and Walkinshaw,M.D. (1994). The molecular replacement solution and X-ray refinement to 2.8 angstrom of a decameric complex of human Cyclophilin A with the immunosuppressive drug Cyclosporin A. Journal Of Molecular Biology *244*, 385-409.

Pratt,W.B. (1998). The Hsp90 based chaperone system: Involvement in signal transduction from a variety of hormone and growth factor receptors. Proceedings Of The Society For Experimental Biology And Medicine *217*, 420-434.

Pratt,W.B., Krishna,P., and Olsen,L.J. (2001). Hsp90 binding immunophilins in plants: the protein movers. Trends in Plant Science *6*, 54-58.

Pratt,W.B. and Toft,D.O. (1997). Steroid receptor interactions with heat shock protein and immunophilin chaperones. Endocrine Reviews *18*, 306-360.

Prodromou,C., Roe,S.M., Obrien,R., Ladbury,J.E., Piper,P.W., and Pearl,L.H. (1997b). Identification and structural characterization of the ATP/ADP binding site in the Hsp90 molecular chaperone. Cell *90*, 65-75.

Prodromou,C., Roe,S.M., Piper,P.W., and Pearl,L.H. (1997a). A molecular clamp in the crystal structure of the N-terminal domain of the yeast Hsp90 chaperone. Nature Structural Biology *4*, 477-482.

Prodromou,C., Siligardi,G., Obrien,R., Woolfson,D.N., Regan,L., Panaretou,B., Ladbury,J.E., Piper,P.W., and Pearl,L.H. (1999). Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR) domain co-chaperones. EMBO Journal *18*, 754-762.

Radanyi,C., Chambraud,B., and Baulieu,E.E. (1994). The ability of the immunophilin FKBP59-HBI to interact with the 90kDa heat shock protein is encoded by its tetratricopeptide repeat domain. Proceedings Of The National Academy Of Sciences Of The United States Of America *91*, 11197-11201.

Radzicka,A., Acheson,S.A., and Wolfenden,R. (1992). *Cis-trans* isomerization at proline - desolvation and its consequences for protein folding. Bioorganic Chemistry *20*, 382-386.

Ramsey,A.J., Russell,L.C., Whitt,S.R., and Chinkers,M. (2000). Overlapping sites of tetratricopeptide repeat protein binding and chaperone activity in heat shock protein 90. Journal Of Biological Chemistry *275*, 17857-17862.

Rassow,J., Mohrs,K., Koidl,S., Barthelmess,I.B., Pfanner,N., and Tropschug,M. (1995). Cyclophilin 20 is involved in mitochondrial protein-folding in co-operation

with molecular chaperones Hsp70 and Hsp60. Molecular And Cellular Biology *15*, 2654-2662.

Ratajczak,T. and Carrello,A. (1996). Cyclophilin 40 (Cyp 40), mapping of its Hsp90 binding domain and evidence that FKBP 52 competes with Cyp 40 for Hsp90 binding. Journal Of Biological Chemistry *271*, 2961-2965.

Ratajczak,T., Carrello,A., Mark,P.J., Warner,B.J., Simpson,R.J., Moritz,R.L., and House,A.K. (1993). The cyclophilin component of the unactivated estrogen receptor contains a tetratricopeptide repeat domain and shares identity with p59 (FKBP59). Journal Of Biological Chemistry *268*, 13187-13192.

Ratajczak,T., Carrello,A., and Minchin,R.F. (1995). Biochemical and calmodulin binding properties of estrogen receptor binding cyclophilin expressed in *Escherichia coli*. Biochemical And Biophysical Research Communications *209*, 117-125.

Ratajczak,T., Hlaing,J., Brockway,M.J., and Hahnel,R. (1990). Isolation of untransformed bovine estrogen receptor without molybdate stabilization. Journal Of Steroid Biochemistry And Molecular Biology *35*, 543-553.

Rebbe,N.F., Ware,J., Bertina,R.M., Modrich,P., and Stafford,D.W. (1987). Nucleotide sequence of a cDNA for a member of the human 90kDa heat-shock protein family. Gene *53*, 235-245.

Reddy,G.R. (1995). Cloning and characterization of a *Plasmodium falciparum* cyclophilin gene that is stage specifically expressed. Molecular And Biochemical Parasitology *73*, 111-121.

Reddy,G.R. and Ingallinera,J. (1995). Gene sequence-analysis of *Plasmodium falciparum* cyclophilin which is stage specifically expressed. Faseb Journal *9*, 145-145.

Renoir,J.M., Radanyi,C., and Baulieu,E.E. (1992). Effect of the immunosuppressants FK506 and rapamycin on progesterone receptor function - the heat-shock protein binding p59 immunophilin between immunological and hormonal activities. Comptes Rendus De L Academie Des Sciences Serie III, Sciences De La Vie-Life Sciences *315*, 421-428.

Reynolds,P.D., Ruan,Y., Smith,D.F., and Scammell,J.G. (1999). Glucocorticoid resistance in the squirrel monkey is associated with overexpression of the immunophilin FKBP 51. Journal Of Clinical Endocrinology And Metabolism *84*, 663-669.

Rice,L.M. and Brunger,A.T. (1999). Crystal structure of the vesicular transport protein Sec17: Implications for SNAP function in SNARE complex disassembly. Molecular Cell *4*, 85-95.

Rinfret,A., Collins,C., Menard,R., and Anderson,S.K. (1994). The N-terminal cyclophilin homologous domain of a 150 kilodalton tumor recognition molecule

exhibits both peptidyl-prolyl *cis-trans-* isomerase and chaperone activities. Biochemistry *33*, 1668-1673.

Rosen,M.K., Standaert,R.F., Galat,A., Nakatsuka,M., and Schreiber,S.L. (1990). Inhibition of FKBP rotamase activity by immunosuppressant FK506 - twisted amide surrogate. Science *248*, 863-866.

Rost,B., Sander,C., and Schneider,R. (1994). Phd - an automatic mail server for protein secondary structure prediction. Computer Applications In The Biosciences *10*, 53-60.

Russell,L.C., Whitt,S.R., Chen,M.S., and Chinkers,M. (1999). Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. Journal Of Biological Chemistry *274*, 20060-20063.

Rutherford,S.L. and Lindquist,S. (1998). Hsp90 as a capacitor for morphological evolution. Nature *396*, 336-342.

Rycyzyn,M.A., Reilly,S.C., O'Malley,K., and Clevenger,C.V. (2000). Role of cyclophilin B in prolactin signal transduction and nuclear retrotranslocation. Molecular Endocrinology *14*, 1175-1186.

Sambrook and Russell. Molecular Cloning, A Laboratory Manual. (2000) Third. Edition, Cold Spring Harbour Laboratory Press.

Scheibel,T., Weikl,T., and Buchner,J. (1998). Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence. Proceedings Of The National Academy Of Sciences Of The United States Of America *95*, 1495-1499.

Scheufler,C., Brinker,A., Bourenkov,G., Pegoraro,S., Moroder,L., Bartunik,H., Hartl,F.U., and Moarefi,I. (2000). Structure of TPR domain-peptide complexes: Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell *101*, 199-210.

Schmitz,G., Schmidt,M., and Feierabend,J. (1996). Characterization of a plastid specific Hsp90 homologue: Identification of a cDNA sequence, phylogenetic descendence and analysis of its mRNA and protein expression. Plant Molecular Biology *30*, 479-492.

Schneider,C., SeppLorenzino,L., Nimmesgern,E., Ouerfelli,O., Danishefsky,S., Rosen,N., and Hartl,F.U. (1996). Pharmacologic shifting of a balance between protein refolding and degradation mediated by Hsp90. Proceedings Of The National Academy Of Sciences Of The United States Of America *93*, 14536-14541.

Scholz,C., Stoller,G., Zarnt,T., Fischer,G., and Schmid,F.X. (1997). Co-operation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. EMBO Journal *16*, 54-58.

Scholz,C., Zarnt,T., Kern,G., Lang,K., Burtscher,H., Fischer,G., and Schmid,F.X. (1996). Autocatalytic folding of the folding catalyst FKBP12. Journal Of Biological Chemistry *271*, 12703-12707.

Schonbrunner,E.R., Mayer,S., Tropschug,M., Fischer,G., Takahashi,N., and Schmid,F.X. (1991). Catalysis of protein folding by cyclophilins from different species. Journal Of Biological Chemistry *266*, 3630-3635.

Schonbrunner,E.R. and Schmid,F.X. (1992). Peptidyl-prolyl cis-trans isomerase improves the efficiency of protein disulfide isomerase as a catalyst of protein folding. Proceedings Of The National Academy Of Sciences Of The United States Of America *89*, 4510-4513.

Sehgal,S.N. (1995). Rapamune (Sirolimus, Rapamycin) - an overview and mechanism of action. Therapeutic Drug Monitoring *17*, 660-665.

Sheldrick, G.M. and Schneider, T.R., (1997) SHELXL: High resolution refinement. Macromolecular Crystallography Pt B, Methods in Enzymology 277, 319-343.

Siegert,R., Leroux,M.R., Scheufler,C., Hartl,F.U., and Moarefi,I. (2000). Structure of the molecular chaperone Prefoldin: Unique interaction of multiple coiled coil tentacles with unfolded proteins. Cell *103*, 621-632.

Siekierka,J.J., Hung,S.H.Y., Poe,M., Lin,C.S., and Sigal,N.H. (1989). A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. Nature *341*, 755-757.

Sikorski,R.S., Boguski,M.S., Goebl,M., and Hieter,P. (1990). A repeating amino acid motif in cdc23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. Cell *60*, 307-317.

Smith,D.B. and Johnson,K.S. (1988). Single step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-S-transferase. Gene *67*, 31-40.

Smith,R.L., Redd,M.J., and Johnson,A.D. (1995). The tetratricopeptide repeats of ssn6 interact with the homeo domain of alpha-2. Genes & Development *9*, 2903-2910.

Solscheid,B. and Tropschug,M. (2000). A novel type of FKBP in the secretory pathway of *Neurospora crassa*. Febs Letters *480*, 118-122.

Song,H.Y., Dunbar,J.D., Zhang,Y.X., Guo,D.Q., and Donner,D.B. (1995). Identification of a protein with homology to Hsp90 that binds the type1 tumor necrosis factor receptor. Journal Of Biological Chemistry *270*, 3574-3581.

Spence,J. and Georgopoulos,C. (1989). Purification and properties of the *Escherichia coli* heat shock protein, HtpG. Journal Of Biological Chemistry *264*, 4398-4403.

Stamnes,M.A., Shieh,B.H., Chuman,L., Harris,G.L., and Zuker,C.S. (1991). The cyclophilin homolog NinaA is a tissue-specific integral membrane protein required for the proper synthesis of a subset of *Drosophila* rhodopsins. Cell *65*, 219-227.

Standaert,R.F., Galat,A., Verdine,G.L., and Schreiber,S.L. (1990). Molecular cloning and overexpression of the human FK506 binding protein FKBP. Nature *346*, 671-674.

Stebbins,C.E., Russo,A.A., Schneider,C., Rosen,N., Hartl,F.U., and Pavletich,N.P. (1997). Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. Cell *89*, 239-250.

Steinmann,B., Bruckner,P., and Supertifurga,A. (1991). Cyclosporin A slows collagen triple helix formation *in vivo* - indirect evidence for a physiological role of peptidyl prolyl *cis trans* isomerase. Journal Of Biological Chemistry *266*, 1299-1303.

Stevens,F.J. and Argon,Y. (1999). Protein folding in the ER. Seminars in Cell & Developmental Biology *10*, 443-454.

Stewart,D.E., Sarkar,A., and Wampler,J.E. (1990). Occurrence and role of *cis* peptide bonds in protein structures. Journal Of Molecular Biology *214*, 253-260.

Sturzenbaum,S.R., Morgan,A.J., and Kille,P. (1999). Characterisation and quantification of earthworm cyclophilins: identification of invariant and heavy metal responsive isoforms. Biochimica et Biophysica Acta-Gene Structure And Expression *1489*, 467-473.

Sykes,K., Gething,M.J., and Sambrook,J. (1993). Proline isomerases function during heat shock. Proceedings Of The National Academy Of Sciences Of The United States Of America *90*, 5853-5857.

Tai,P.K.K., Albers,M.W., Chang,H., Faber,L.E., and Schreiber,S.L. (1992). Association of a 59 kilodalton immunophilin with the glucocorticoid receptor complex. Science *256*, 1315-1318.

Tai,P.K.K., Chang,H., Albers,M.W., Schreiber,S.L., Toft,D.O., and Faber,L.E. (1993). P59 (FK506 binding protein 59) interaction with heat-shock proteins is highly conserved and may involve proteins other than steroid-receptors. Biochemistry *32*, 8842-8847.

Takahashi,N., Hayano,T., and Suzuki,M. (1989). Peptidyl prolyl *cis-trans* isomerase is the Cyclosporin A binding protein cyclophilin. Nature *337*, 473-475.

Tanaka,H., Kuroda,A., Marusawa,H., Hatanaka,H., Kino,T., Goto,T., Hashimoto,M., and Taga,T. (1987). Structure of FK506 - A novel immunosuppressant isolated from *Streptomyces*. Journal Of The American Chemical Society *109*, 5031-5033.

Tang,N.M., Ho,C.Y., and Katze,M.G. (1996). The 58kDa cellular inhibitor of the double stranded RNA dependent protein kinase requires the tetratricopeptide repeat 6

and DnaJ motifs to stimulate protein synthesis *in vivo*. Journal Of Biological Chemistry *271*, 28660-28666.

Taylor,P., Dornan,J., Carrello,A., Minchin,R.F., Ratajczak,T., and Walkinshaw,M.D. (2001). Two structures of cyclophilin 40: Folding and fidelity in the TPR domains. Structure *9*, 431-438.

Taylor,P., Husi,H., Kontopidis,G., and Walkinshaw,M.D. (1997). Structures of cyclophilin-ligand complexes. Progress In Biophysics & Molecular Biology *67*, 155-181.

Taylor,P., Page,A.P., Kontopidis,G., Husi,H., and Walkinshaw,M.D. (1998). The X-ray structure of a divergent cyclophilin from the nematode parasite *Brugia malayi*. Febs Letters *425*, 361-366.

Thali,M., Bukovsky,A., Kondo,E., Rosenwirth,B., Walsh,C.T., Sodroski,J., and Gottlinger,H.G. (1994). Functional association of Cyclophilin A with HIV-1 virions. Nature *372*, 363-365.

Timerman,A.P., Ogunbumni,E., Freund,E., Wiederrecht,G., Marks,A.R., and Fleischer,S. (1993). The calcium release channel of sarcoplasmic reticulum is modulated by FK506 binding protein - dissociation and reconstitution of FKBP 12 to the calcium rlease channel of skeletal muscle sarcoplasmic reticulum. Journal Of Biological Chemistry *268*, 22992-22999.

Tropschug,M., Wachter,E., Mayer,S., Schonbrunner,E.R., and Schmid,F.X. (1990). Isolation and sequence of an FK506 binding protein from *Neurospora crassa* which catalyzes protein folding. Nature *346*, 674-677.

Vagin,A. and Teplyakov,A. (1997). MOLREP: an automated program for molecular replacement. Journal Of Applied Crystallography *30*, 1022-1025.

Vanduyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L., and Clardy,J. (1991a). Atomic-structure of FKBP-FK506, an immunophilin- immunosuppressant complex. Science *252*, 839-842.

Vanduyne,G.D., Standaert,R.F., Schreiber,S.L., and Clardy,J. (1991b). Atomic structure of the rapamycin human immunophilin FKBP 12 complex. Journal Of The American Chemical Society *113*, 7433-7434.

Vetter,I.R., Nowak,C., Nishimoto,T., Kuhlmann,J., and Wittinghofer,A. (1999). Structure of a Ran binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. Nature *398*, 39-46.

Vijayalakshmi,J., Mukhergee,M.K., Graumann,J., Jakob,U., and Saper,M.A. (2001). The 2.2 angstrom crystal structure of Hsp33: A heat shock protein with redox regulated chaperone activity. Structure *9*, 367-375.

Vittorioso,P., Cowling,R., Faure,J.D., Caboche,M., and Bellini,C. (1998). Mutation in the *Arabidopsis* PASTICCINO1 gene, which encodes a new FK506 binding

protein like protein, has a dramatic effect on plant development. Molecular And Cellular Biology *18*, 3034-3043.

Walker,J.E., Saraste,M., Runswick,M.J., and Gay,N.J. (1982). Distantly related sequences in the alpha subunits and beta subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. EMBO Journal *1*, 945-951.

Warth,R., Briand,P.A., and Picard,D. (1997). Functional analysis of the yeast 40 kDa cyclophilin Cyp 40 and its role for viability and steroid receptor regulation. Biological Chemistry *378*, 381-391.

Wearsch,P.A., Voglino,L., and Nicchitta,C.V. (1998). Structural transitions accompanying the activation of peptide binding to the endoplasmic reticulum Hsp90 chaperone Grp94. Biochemistry *37*, 5709-5719.

Weisman,R., Creanor,J., and Fantes,P. (1996). A multicopy suppressor of a cell-cycle defect in *Schizosaccharomyces pombe* encodes a heat shock inducible 40kDa cyclophilin like protein. EMBO Journal *15*, 447-456.

Welch,W.J. and Feramisco,J.R. (1982). Purification of the major mammalian heat-shock proteins. Journal Of Biological Chemistry *257*, 4949-4959.

White,J. (1988). The Anatomy. In *The Nematode Caenohabditis elegans*, (Wood, W.B.E., ed. Cold Spring Harbor Laboratory), pp. 81-123.

Whitesell,L. and Cook,P. (1996). Stable and specific binding of heat shock protein 90 by geldanamycin disrupts glucocorticoid receptor function in intact cells. Molecular Endocrinology *10*, 705-712.

Whitesell,L., Mimnaugh,E.G., Decosta,B., Myers,C.E., and Neckers,L.M. (1994). Inhibition of heat shock protein Hsp90-pp60(v-src) heteroprotein complex-formation by benzoquinone ansamycins - essential role for stress proteins in oncogenic transformation. Proceedings Of The National Academy Of Sciences Of The United States Of America *91*, 8324-8328.

Whitesell,L., Sutphin,P., An,W.G., Schulte,T., Blagosklonny,M.V., and Neckers,L. (1997). Geldanamycin stimulated destabilization of mutated p53 is mediated by the proteasome *in vivo*. Oncogene *14*, 2809-2816.

Whitesell,L., Sutphin,P.D., Pulcini,E.J., Martinez,J.D., and Cook,P.H. (1998). The physical association of multiple molecular chaperone proteins with mutant p53 is altered by geldanamycin, an Hsp90- binding agent. Molecular And Cellular Biology *18*, 1517-1524.

Widmer,A. (1997). WITNOTP: A computer program for molecular modelling. Novartis AG, Basel.

Wiederrecht,G., Brizuela,L., Elliston,K., Sigal,N.H., and Siekierka,J.J. (1991). Fkb1 encodes a nonessential FK506 binding protein in *Saccharomyces cerevisiae* and

contains regions suggesting homology to the cyclophilins. Proceedings Of The National Academy Of Sciences Of The United States Of America *88*, 1029-1033.

Wilkins M R and Gasteiger E, Baroch A Sanchez J-C Williams KL Appel RD Hochstrasser DF. Protein identification and analysis tools in the ExPASy Server. Ed A.J.Link. 2-D Proteome Analysis Protocols. 1998. New York, Humana Press. Ref Type: Generic

Wilmot,C.M. and Thornton,J.M. (1988). Analysis and prediction of the different types of beta turn in proteins. Journal Of Molecular Biology *203*, 221-232.

Wilson,K.P., Yamashita,M.M., Sintchak,M.D., Rotstein,S.H., Murcko,M.A., Boger,J., Thomson,J.A., Fitzgibbon,M.J., Black,J.R., and Navia,M.A. (1995). Comparative X-ray structures of the major binding-protein for the immunosuppressant FK506 (tacrolimus) in unliganded form and in complex with FK506 and rapamycin. Acta Crystallographica Section D-Biological Crystallography *51*, 511-521.

Wu,S.Y., Dornan,J., Kontopidis,G., Taylor,P., and Walkinshaw,M.D. (2001). The first direct determination of a ligand binding constant in protein crystals. Angewandte Chemie-International Edition *40*, 582-586.

Yem,A.W., Tomasselli,A.G., Heinrikson,R.L., Zurcherneely,H., Ruff,V.A., Johnson,R.A., and Deibel,M.R. (1992). The hsp56 component of steroid receptor complexes binds to immobilized FK506 and shows homology to FKBP-12 and FKBP-13. Journal Of Biological Chemistry *267*, 2868-2871.

Yokoyama,N., Hayashi,N., Seki,T., Pante,N., Ohba,T., Nishii,K., Kuma,K., Hayashida,T., Miyata,T., Aebi,U., Fukui,M., and Nishimoto,T. (1995). A giant nucleopore protein that binds Ran/tc4. Nature *376*, 184-188.

Young,J.C., Obermann,W.M.J., and Hartl,F.U. (1998). Specific binding of tetratricopeptide repeat proteins to the C terminal 12kDa domain of Hsp90. Journal Of Biological Chemistry *273*, 18007-18010.

Young,J.C., Schneider,C., and Hartl,F.U. (1997). *In vitro* evidence that Hsp90 contains two independent chaperone sites. Febs Letters *418*, 139-143.

Yue,L., Karr,T.L., Nathan,D.F., Swift,H., Srinivasan,S., and Lindquist,S. (1999). Genetic analysis of viable Hsp90 alleles reveals a critical role in *Drosophila* spermatogenesis. Genetics *151*, 1065-1079.

Zachariae,W. and Nasmyth,K. (1996). TPR proteins required for anaphase progression mediate ubiquitination of mitotic b type cyclins in yeast. Molecular Biology Of The Cell *7*, 791-801.

Zhao,Y.D. and Ke,H.M. (1996). Mechanistic implication of crystal structures of the cyclophilin dipeptide complexes. Biochemistry *35*, 7362-7368.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 49, Issue of December 3, pp. 34877–34883, 1999
Printed in U.S.A.

Biochemical and Structural Characterization of a Divergent Loop Cyclophilin from *Caenorhabditis elegans**

(Received for publication, June 28, 1999)

Jacqueline Dornan, Antony P. Page‡, Paul Taylor, Su-ying Wu, Alan D. Winter‡, Holger Husi, and Malcolm D. Walkinshaw§

From the Structural Biochemistry Group, Institute of Cell and Molecular Biology, The University of Edinburgh, Michael Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland and ‡Wellcome Centre for Molecular Parasitology, The Anderson College, The University of Glasgow, 56 Dumbarton Road, Glasgow G11 6NU, Scotland

Cyclophilin 3 (CYP-3) is one of the most abundantly expressed cyclophilin isoforms in the free living nematode *Caenorhabditis elegans*. The detailed post-embryonic expression pattern of the *cyp-3* transcript is unusual, peaking during early larval development. The spatial expression pattern was examined via reporter gene analysis demonstrating that the *cyp*-3 transcript is exclusively expressed in the single anterior excretory cell. Recombinant cyclophilin 3 has been purified, crystallized and solved to a resolution of 1.8 Å. The peptidyl-prolyl isomerase activity of CYP-3 has been characterized against the substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide, and gives a k_{cat}/K_m value of 2.4×10^6 M^{-1} s^{-1}. The immunosuppressive drug cyclosporin A binds and inhibits CYP-3 with an IC_{50} value of 16 nM, comparable with the range of values found for human cyclophilin A. The x-ray structure shows that the overall fold and active site geometry is similar to other cyclophilin structures. There are however a number of distinctive features, and we use this structure and amino acid sequence alignment analysis to identify a subgroup of "divergent-loop cyclophilins". This subgroup has a number of uniquely conserved features: an additional loop between residues 48 and 54 (KSGKPLH); two cysteine residues (Cys^{40} and Cys^{168}) that are in close proximity but remain in the unoxidized form, and two other conserved residues, His^{54} and Glu^{83}. We suggest that these features are functionally important for the role played by this class of cyclophilins during cellular responses to stress caused by changes in the redox environment or by up-regulation of cellular activity. This study represents a detailed biological, biochemical, and structural characterization of a single cyclophilin isoform in the model organism *Caenorhabditis elegans*.

Cyclophilins are ubiquitous, highly conserved proteins found in diverse prokaryotic and eukaryotic organisms (for review, see Ref. 1). The first protein to be discovered was the abundant cytosolic protein cyclophilin A (CypA),[1] isolated by its high affinity binding to the natural fungal cyclic undecapeptide cyclosporin A (2, 3). An 18-kDa protein independently isolated (4) showed peptidyl-prolyl isomerase (PPIase) activity and was subsequently shown to be identical to CypA. *In vitro,* cyclophilins accelerate the rate-limiting *cis*/*trans* PPIase steps in the folding pathway of peptide and protein substrates (5, 6), especially in proline-rich proteins (7, 8). It has also been suggested that cyclophilins affect the dynamics of protein-protein interactions *in vitro* by preventing aggregation of early folding intermediates of protein substrates (9). *In vivo,* several lines of evidence point to a role for cyclophilins in protein folding, transport, and assembly (1, 10).

The best characterized example of cyclophilin function *in vivo* is the 26-kDa photoreceptor cyclophilin *ninaA* from *Drosophila melanogaster*. The *ninaA* gene product has been shown to be essential for the proper folding and intracellular transport of isoforms of the photopigment rhodopsin, RH1 and RH2, from the rough endoplasmic reticulum through the cytoplasm to the cell surface (11, 12). Evidence for other cyclophilin functions is growing and includes, for single domain cyclophilins, the requirement for functional PPIase activity of CypA to produce functional expression of homo-oligomeric α7-neuronal nicotinic and type 3 serotonin receptors in *Xenopus* oocytes (13). PPIase activity of the cyclophilin from *Bacillus subtilis* has also been shown to be necessary to allow growth of the organism under starvation conditions (14). In cooperation with Hsp 70 and Hsp 60, cyclophilins are a component of the mitochondrial protein-folding machinery in *Neurospora crassa* (15). There is evidence that the larger cyclophilin 40 proteins induce conformational changes in the transcriptional factor c-Myb which allows the C-terminal negative regulatory domain of c-Myb to interact with and inhibit its N-terminal DNA binding domain (16). Duina *et al.* (17) also suggest a requirement for a Cyp 40-type cyclophilin (Cpr 7) in the negative regulation of the heat shock response under both stress and nonstress conditions in *Saccharomyces cerevisiae*.

The cloning and characterization of eleven cyclophilin genes from *C. elegans* (18) provides an excellent opportunity to study the various biological roles played by representatives of all the cyclophilin subfamilies present within this system. The immunosuppressant drug cyclosporin A has been shown to have an effect on parasitic helminth worms, causing structural tegumental damage. A possible mechanism for this antiparasitic effect is the inhibition of proper cyclophilin activity (19–21) resulting in production and accumulation of misfolded protein. Detailed structural information on the nematode cyclophilins may therefore provide useful targets for the design of new families of antiparasitic drugs.

In this study we have fully characterized both the x-ray structure and enzymatic activity of CYP-3 and also monitored its expression pattern throughout the post-embryonic life cycle of *C. elegans*. The expression pattern analysis of CYP-3, when considered together with emerging structural and sequence

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. E-mail: m.walkinshaw@ed.ac.uk.

[1] The abbreviations used are: CypA, cyclophilin A; PPIase, peptidyl-prolyl isomerase; PCR, polymerase chain reaction; RT-PCR, reverse transcriptase-PCR; MPEG, methoxypolyethylene glycol; CsA, cyclosporin A.

data, provides an insight into the possible roles played by cyclophilins in multicellular organisms.

MATERIALS AND METHODS

Construction of C. elegans CYP-3 Reporter Gene Fusions—The *cyp-3* reporter gene recombinant plasmids were constructed using the *C. elegans lacZ* promotorless reporter gene expression vector pPD95.03 (a kind gift from Andy Fire and co-workers, Carnegie Institute). The plasmid encodes a nuclear localization signal amino-terminal to the *lacZ* gene and an N-terminal multi-intron intervening sequence. The potential upstream regulatory region of *cyp-3* was identified from an unfinished yeast artificial chromosome clone (Y75B12, GenBank™ accession number AL022301), identified on chromosome *V* by the *C. elegans* genome project. The translationally "in-frame" 1026-base pair 5′ untranslated region of the *cyp-3* gene was amplified from genomic DNA by polymerase chain reaction (PCR) with *Taq* polymerase (Applied Biotechnologies). The PCR primers contained an artificial restriction site (underlined) to allow directional subcloning into the multiple cloning site of the vector pPD95.03; cyp3pfxba 5′-gcgtctagaagaagtccaaattactgtcaac-3′ (sense, *Xba*I); cyp3prbam 5′-gcgggatcctcattgttgaggcaagaataggga-3′ (antisense, *Bam*HI). This construct permitted an in-frame translational fusion to the *lacZ* with the first two amino acids of CYP-3. PCR products were sequentially digested with *Xba*I and *Bam*HI, gel purified, and ligated into similarly digested vector pPD95.03. Miniprep DNA was prepared using a Qiagen miniprep kit, and sequencing was performed to confirm the identity and the translational context of the insert.

DNA Transformation, Fixation, and β-Galactosidase Staining of C. elegans—Transformation of adult hermaphrodite *C. elegans* was performed by microinjection of plasmid DNA into the syncytial gonad and was carried out using the standard methods as described previously (22). Transformed nematodes were fixed and stained for β-galactosidase activity as described previously (22).

Semi-quantitative Reverse Transcriptase (RT)-PCR—The semi-quantitative RT-PCR methods including the generation of synchronous nematode cultures for staged mRNA and subsequent cDNA are described in detail elsewhere (23). The staged cDNA samples were supplied by Iain Johnstone (Wellcome Centre for Molecular Parasitology, The University of Glasgow). The gene combinations (*cyp-3* and the control gene *ama-1,* which encodes the large subunit of RNA polymerase II) were amplified from the staged cDNA samples representing pooled mRNA samples corresponding to the different post-embryonic developmental stages, using conditions which allowed reactants to remain in excess (94 °C for 30 s, 58 °C for 60 s, and 72 °C for 90 s for 25 cycles in 25 μl of final volume). Primers were designed to span introns so as to distinguish between cDNA and possible genomic DNA signals, and combinations used were as follows: *cyp-3*, cyp3F/Nde and cyp3R/*Eco*RI *ama-1*, ama1f 5′-ttccaagcgccgctgcgcattgtctc-3′ and ama1r 5′-cagaatttccagcactcgaggagcgga-3′. The PCR reactions were electrophoresed on a 2% agarose gel and Southern blotted; blots were then probed with corresponding gene-specific (*cyp-3*, and the control gene *ama-1*) dAT-^{32}P randomly labeled inserts, which were generated by PCR using the above conditions and purified by gel extraction (QiaQuick gel extraction kit). After autoradiography, the bands corresponding to the respective genes were excised and counted by scintillation. The relative abundance of the test genes were reproducibly determined by comparing their signal with that of the control gene *ama-1*.

pET Expression of CYP-3—cyp-3 was cloned from cDNA using a RT-PCR method as described previously (18). Briefly, poly(A)$^+$ mRNA was obtained using the Micro-Fast Track Kit (Invitrogen), and cDNA was synthesized using a cDNA synthesis kit (Amersham Pharmacia Biotech, Little Chalfont, UK). RT-PCR conditions were as follows: 10 pmol of each primer (cyp3F/Nde and cyp3R/*Eco*RI) 5 mM $MgCl_2$, 6.7 mM 2-mercaptoethanol, 4.4 μM EDTA (pH 8), 1 mM of each dNTP, 113 μg/ml nonacetylated bovine serum albumin, 2 μl of mixed stage cDNA and 1 μl of *Taq* polymerase (Advanced Biotechnologies), cycled 20× at 94 °C for 30 s, 60 °C for 1 min, and 72 °C for 1 min in 100 μl. The primers included engineered restriction sites (*Nde*I and *Eco*RI) that, after restriction digestion, allowed directional subcloning into similarly digested pET-5a expression vector (Promega). Ligation and transformation protocols of the Promega pET-5 technical manual were followed. Plasmid DNA was isolated (Plasmid Kit, Qiagen), and the insert was sequenced in both directions by thermal cycle sequencing on an Applied Biosystems automated sequencer to validate that no mutations had been generated by the RT-PCR procedure. The plasmid DNA was then transformed into competent BL21(DE3 pLysS) cells (Promega) for subsequent recombinant protein expression.

TABLE I
Summary of the final refinement statistics used in the structure determination of cyclophilin 3

Summary of refinement statistics	Native
Resolution range (Å)	10–1.8
$R_{final\ (all\ data)}$ ($F > 4\ \sigma$)	20.81% (19.69%)
R_{free} ($F > 4\ \sigma$)	28.47% (26.61%)
Root mean square deviations from ideal geometry	
Bond lengths (Å)	0.031
Angle distance (Å)	0.026
Mean B-factor (Å^2)	30.37

Purification of Recombinant Cyclophilin 3—Recombinant CYP-3 was expressed following a 4 h induction with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside. Cell pellets were frozen at −20 °C, defrosted on ice, and solubilized in lysis buffer (10% w/v) containing 50 mM HEPES, 5 mM benzamidine, 5 mM EDTA, 5 mM 2-mercaptoethanol, pH 7.5. Lysozyme (ICN/FLOW) was added to 0.1% w/v and incubated on ice for 1 h. An equal volume of deionized water at 4 °C was added, and the cell suspension mixed. Resulting supernatant was frozen at −20 °C, defrosted slowly on ice and the pH adjusted to 6.8. Protein extract was purified to homogeneity using two purification steps: SP-Sepharose cation exchange chromatography with the following buffers: Buffer A, 50 mM HEPES, 5 mM EDTA, 5 mM 2-mercaptoethanol, pH 6.8, and Buffer B, Buffer A + 0.5 M NaCl (pH 6.8). Column fractions were tested on 15% SDS-polyacrylamide electrophoresis gels and fractions containing the protein of interest were dialyzed overnight against Buffer A. Protein solution was filtered through a 0.2 μm filter and applied to Resource S resin (Amersham Pharmacia Biotech). CYP-3 was eluted from this resin with the following buffers: Buffer A, 20 mM HEPES, pH 6.8, Buffer B, Buffer A + 0.5 M NaCl.

Fractions containing CYP-3 were pooled and concentrated to 20 mg/ml as estimated by the Bradford assay. Dynamic light scattering showed the protein was essentially monodisperse and stable with an estimated molecular mass of 20 kDa (results not shown).

PPIase Assay—Purified recombinant CYP-3 was assayed for PPIase activity essentially as described by Fischer *et al.* (4), with the suggested substrate solvent application as described by Kofron *et al.* (24). This assay determines the rate of conversion of *cis* to *trans* of a proline-containing peptide, working on the principle that chymotrypsin will cleave the peptide only when it is in the *trans* conformation, resulting in the release of chromogenic dye. The peptide substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Bachem) was dissolved in trifluroethanol with 470 mM LiCl to give 100 mM stock solution, which was diluted to 4 mM immediately before use. The experimental setup was as follows: CYP-3 was diluted into 50 mM HEPES, 86 mM NaCl, pH 8.0 (PPIase Buffer), to give a 2,000 nM stock. In a 1-ml cuvette, 10 μl of protein (20 nM final concentration) was added to 860 μl of PPIase buffer and incubated on ice for 1 h. Chymotrypsin (100 μl of a 6 mg/ml solution) was added to the cuvette followed by transfer of the cuvette to a thermostatted Peltier cell holder by which the reaction temperature was maintained at 5 °C. The sample was allowed to equilibrate until a steady baseline was established. The reaction was started by the addition of 30 μl (120 μM) of peptide substrate followed by rapid mixing. The absorbance change at 400 nm due to the release of *p*-nitroaniline was recorded using a Perkin-Elmer Lambda 20 spectrophotometer with data collected every 0.1 s over a 2-min time period. A number of protein concentrations were tested using the above method. Analysis of the data was carried out using a double reciprocal Lineweaver-Burke plot, which gives an intercept at $1/V_{max}$ on the *y* axis, and $1/[S] = -1/K_m$ on the *x* axis. The slope of the line is K_m/V_{max}.

Cyclosporin A inhibition assays were carried out essentially as described above except that the samples were preincubated for a minimum of 60 min with varying amounts of CsA (kindly provided by Novartis AG).

Crystallization of CYP-3—Two crystal forms were grown from the protein isolated above using the standard hanging drop vapor diffusion method. Crystal forms were tetragonal bipyramid and thin plates.

Tetragonal Form—4-μl hanging drops consisted of 50 mM sodium citrate, pH 5.6, 15.5% w/v methoxypolyethylene glycol 5000 (MPEG 5000). Protein concentration was 3.5 mg/ml. The drop was suspended over the well containing 100 mM sodium citrate, 31% MPEG 5000 at 18 °C. Conditions for producing tetragonal crystals were refined, crystals grew in a narrow concentration range of MPEG 5000 (28–32%) with the optimum pH for crystal growth being 5.6.

Thin Plate Form—4-μl hanging drops consisted of 50 mM citric acid/

```
CYP3_CAEEL  ----------MSRSKVFFDITIGGKASGRIVMELYDDVVPKTAGNFRALC 40
CYPH_TCRUZ  MRKINVPSKGERRRGIFVEVSRNPKVFFDISIDNKPSGRIKMELYADTVPKTAENFRALC 60
CYP1_ARATH  ----------MAFPKVYFDMTIDQQPAGRIVMELYTDKTPRTAENFRALC 40
CYPH_MAIZE  ----------MANPRVFFDMTVGGAPAGRIVMELYANEVPKTAENFRALC 40
CYPH_SYNEC  ----------MMSKVFFDITIGSDTAGRIVMELFDEVTPKTAENFRALC 39
CYP4_BOVIN  --------MSHPSPQAKPSNPSNPRVFFDVDIGGERVGRIVLELFADIVPKTAENFRALC 52
CYP1_BRMAL  ----------MSKKDRRRVFLDVTIDGMLAGRIVMELYNDIAPRTCNNFLMLC 43
CYPH_ONCHO  ----------MSKKDRRRVFFDVTIDGNLAGRIVMELYNDIAPRTCNNFLMLC 43
CYPH_HUMAN  ----------MVNPTVFFDIAVDGEPLGRVSFELFADKVPKTAENFRALS 40
                        *:;*: :.:   **: :**: : .:*. ** *.

CYP3_CAEEL  TGENGIGK-SGKPLHFKGSKFHRIIPNFMIQGGDFTRGNGTGGESIYGEKFPDENF---K 96
CYPH_TCRUZ  TGEKGRGK-SGKPLHYKGCVFHRVIPGFMIQGGDITRGNGTGGESIYGMNFRDESFSGKA 119
CYP1_ARATH  TGEKGVGG-TGKPLHFKGSKFHRVIPNFMCQGGDFTAGNGTGGESIYGSKFEDENF---E 96
CYPH_MAIZE  TGEKGVGK-SGKPLHYKGSTFHRVIPEFMCQGGDFTRGNGTGGESIYGEKFPDEKF---V 96
CYPH_SYNEC  TGEKGVGKA-GKPLHFKGSSFHRVITDFMAQGGDFTRGNGTGGESIYGEKFADENF---Q 95
CYP4_BOVIN  TGEKGIGPTTGKPLHFKGCPFHRIIKKFMIQGGDFSNQNGTGGESIYGEKFEDENF---H 109
CYP1_BRMAL  TGMAGTGKI SGKPLHYKGSTFHRVIKNFMIQGGDFTKGDGTGGESIYGGMFDDEEF---V 100
CYPH_ONCHO  TGMAGIGKI SGKPLHYKGSTFHRVIKNFMIQGGDFTKGDGTGGESIYGGMFDDEEF---V 100
CYPH_HUMAN  TGEKGFG--------YKGSCFHRIIPGFMCQGGDFTRHNGTGGKSIYGEKFEDENF---I 89
            **  * *         :**. ***:*  ** ****:: :****:**** * **.*

CYP3_CAEEL  EKHTGPGVLSMANAGPNTNGSQFFLCTVKTEWLDGKHVVFGRVVEGLDVVKAVESNGSQS 156
CYPH_TCRUZ  GKHTGVGCLSMANAGPNTNGSQFFICTANTPWLDGKHVVFGRVTEGIDVVRRVERLGSES 179
CYP1_ARATH  RKHTGPGILSMANAGANTNGSQFFICTVKTDWLDGKHVVFGQVVEGLDVVKAIEKVGSSS 156
CYPH_MAIZE  DKQDAPGILSMANAGPNTNGSQFFICTIATPWLDGKHVVFGRVIEGMDVVKAIEKVGTRN 156
CYPH_SYNEC  LKHDRPGLLSMANAGPNTNGSQFFLTFVPCPWLDGKHVVFGEVVEGLEILEQLEANGSQS 155
CYP4_BOVIN  YKHDKEGLLSMANAGSNTNGSQFFITTVPTPWLDGKHVVFGQVIKGMGVAKILENVEVKG 169
CYP1_BRMAL  MKHDEPFVVSMANKGPNTNGSQFFITTTPAPHLNNIHVVFGKVVSGQEVVTKIEYLKTNS 160
CYPH_ONCHO  MKHDEPFLVSMANKGPNTNGSQFFITTTPAPHLNNIHVVFGKVVSGQKVVTKIECLKTNS 160
CYPH_HUMAN  LKHTGPGILSMANAGPNTNGSQFFICTAKTEWLDGKHVVFGKVKEGMNIVEAMERFGSRN 149
             *:    :**** *.********:  .   *:  *****.* .*  :   :*    .

CYP3_CAEEL  -GKPVKDCMIADCGQLKA---------------------------------- 173
CYPH_TCRUZ  -GKTRGRIIIADCGEVVVPEPK*---------------------------- 200
CYP1_ARATH  -GKPTKPVVVADCGQLS----------------------------------- 172
CYPH_MAIZE  -GSTSKVVKVADCGQLS----------------------------------- 172
CYPH_SYNEC  -GQTKQAIVISDCGEIK----------------------------------- 171
CYP4_BOVIN  -EKPAKLCVIAECGELKEGDDM*------------------------------ 190
CYP1_BRMAL  KNRPLADVVILNCGELVRRK*------------------------------- 180
CYPH_ONCHO  KNRPLADVIILNCGELVRRK*------------------------------- 180
CYPH_HUMAN  -GKTSKKITIADCGQLE----------------------------------- 165
              .    : :**::
```

FIG. 1. **BLAST and FASTA searches were carried out using the sequence GIGKSGKPLHFKG as the target, producing a combined total of 50 hits**. Using cyclophilin as the query and all available data bases at the EBI produced 196 hits. Manual inspection of these cyclophilin sequences revealed additional insert containing sequences that also showed homology to the target sequence and were therefore included in the analysis. Sequences (or cyclophilin domains only) were retrieved in PIR format and aligned using CLUSTAL W. Alignments were carried out both with and without human CypA. Manual inspection of the two multiple alignments produced allowed identification of residues conserved exclusively between divergent-loop containing cyclophilins as opposed to those conserved between all members of the family. Representative examples covering a diverse range of species are shown aligned against human CypA (CYPH HUMAN), included as the cyclophilin archetype. Sequence alignment of the cyclophilin domains from *C. elegans* Cyp 3, accession number P52011 (CYP3 CAEEL); *Trypanosoma cruzi,* accession number O76990 (CYPH TCRUZ); cyclophilin 1 from *Arabidopsis thaliana,* accession number P34790 (CYP1 ARATH); *Zea mays*, P21569 (CYPH MAIZE); *Synechocystis sp.*, accession number P73789 (CYPH SYNEC); bovine Cyp 40, accession number P26882 (CYP4 BOVIN); *B. malayi*, accession number Q27450 (CYPH BRMAL); Cyp 1 from *Onchocerca volvulus,* accession number AAD09564 (CYPH ONCHO); and human CypA, accession number P05092 (CYPH HUMAN). The *asterisk* (*) inserted at the end of an amino acid sequence indicates the sequence is incomplete, with the cyclophilin domain only being used to produce the alignment seen. The sequences were aligned by CLUSTAL W (50) with *dashes* introduced for maximum alignment. Inserted residues forming the divergent loop seen in all the aligned sequences with the exception of human CypA are in *bold* and *underlined*. Cysteine, histidine, and glutamate residues conserved within this subgroup of cyclophilins are also shown in *underlined bold*. Residues important in conferring CsA sensitivity are in *bold italics*. Numbering refers to the amino acid sequence of CYP-3.

sodium citrate, pH 6.0, 13% MPEG 5000. Protein concentration was 21 mg/ml. The drop was suspended over a well containing 100 mM citric acid/sodium citrate, pH 6.0, and 26% MPEG 5000 at 4 °C. X-ray data were also collected from the thin plate crystal. These crystals were orthorhombic and diffracted to 1.9 Å in one direction, with unit cell dimensions of a = 58.9 Å, b = 67.5 Å, c = 76.4 Å. No further studies were carried out using this crystal form as they were severely anisotropic.

Structure Determination of the Tetragonal Form of CYP-3—Crystals were transferred from the hanging drop using an appropriately sized cryo-loop into a cryoprotectant solution containing mother liquor plus 20% glycerol. Crystals were removed from the cryoprotectant solution after 5 s and mounted in a cryo-loop (Hampton Research) and flash frozen in liquid nitrogen. X-ray data were collected at 100 K (Oxford cryosystems) using a MAR image plate and processed using DENZO (25). The structure was solved by molecular replacement using the program AMORE (26) with the human cyclophilin A structure as the model. Position and B-factor refinement were performed using SHELX97 (27), water molecules were added using SHELXWAT (27).

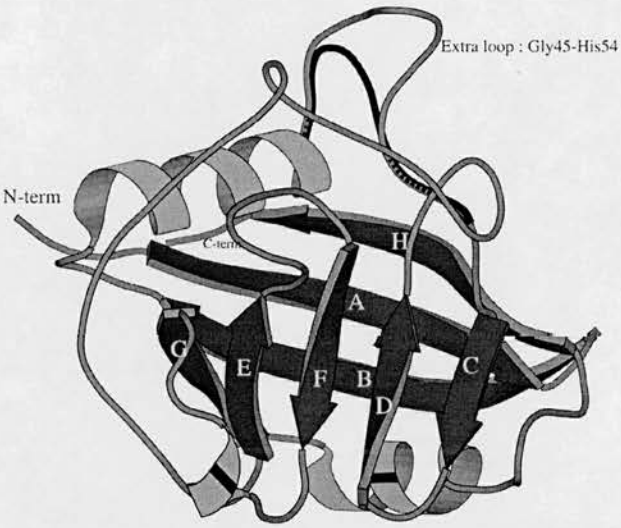


FIG. 2. **Molscript (51) overlay of the stuctures of *C. elegans* cyclophilin 3 and human cyclophilin A**. The eight labeled strands are shown in *purple*, the two helices are shown in *green*, and ceCYP3 loop structure is shown in *orange*. The human CypA loop (in *blue*) contrasts with the shape adopted by the divergent loop, which is formed mainly by 7 additional residues (^{48}KSGKPLH^{54}) in the CYP-3 sequence.

RESULTS

Cyclophilin-3 3D Structure

Crystals belong to space group $P4_12_12$ with unit cell dimensions $a = b = 60.60$ Å, $c = 123.12$ Å. There is one molecule in the asymmetric unit, with $V_m = 3.118$ Å^3/Da and 60.6% solvent. 145,422 observations were measured to give 22,053 unique data, 14,736 with $I > 3\sigma$ (I). The merging R-factor is 8.3%, and the data is 97% complete to a resolution of 1.8 Å. Results of the final refinement are summarized in Table I. All of the residues are well defined in the electron density map, the exceptions being Met^1 and Ser^2 in the N-terminal region, Gly^{87} and Glu^{88} in the loop region and Lys^{172} in the C-terminal region.

CYP-3 and CypA are 67% identical (77% similar) at the amino acid level (18). Fig. 1 shows the alignment of the cyclophilin domains from a number of cyclophilins against human CypA. The overall architecture of CYP-3 is similar to that of known x-ray and NMR cyclophilin structures (28, 29) with an anti-parallel 8-stranded β-barrel capped by two α-helices. The CYP-3 structure is very highly conserved when compared with the structure of human cyclophilin A, the best characterized member of the family (30, 31) (Fig. 2). The root mean square deviation is 0.55 Å for the 620 corresponding C, N, Cα, and O atoms from amino acid residues 2 to 44 and 55 to 171 in CYP-3 and residues 2 to 44 and 48 to 164 in CypA.

The active site of CYP-3 is also very similar to that found in human CypA, all 13 residues important in binding to cyclosporin A being conserved (32) (Fig. 3). The only significant structural difference seen is in the conformation adopted by Arg^{62}, which places the corresponding side chain atoms 0.7 Å apart. This fits with the similar enzymatic activity and inhibition by CsA as is found for human CypA. The very close structural similarity of the active site region also suggests that the possible natural cellular substrates are also similar between *C. elegans* and mammals.

The Divergent loop—The most distinctive structural feature distinguishing CypA and CYP-3 is the additional insert of 7 residues ^{48}KSGKPLH^{54}, which forms a protruding loop located above the active site of CYP-3 (Fig. 4). This feature has also been observed in the Bm-CYP-1 cyclophilin structure (33, 34)

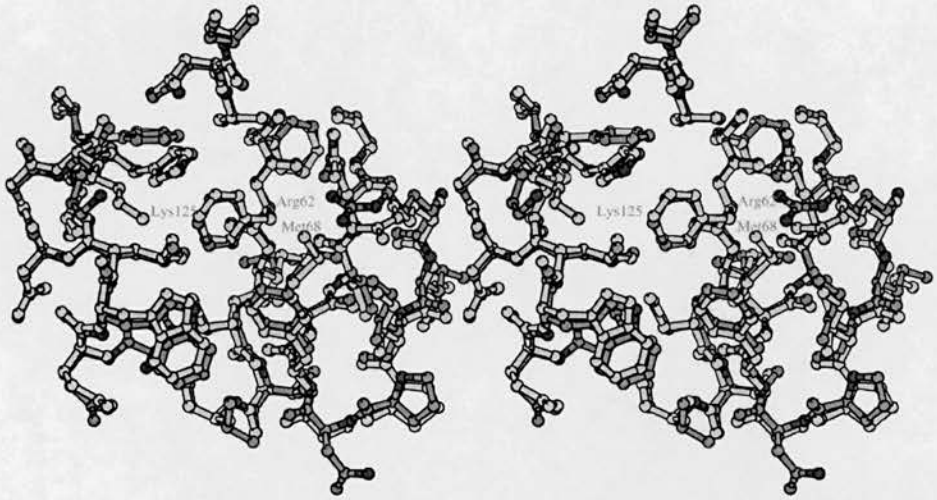


FIG. 3. **Stereo diagram showing the active sites of CYP-3 and human cyclophilin A**. This clearly shows the complete conservation of all 13 residues important in the binding of CsA, with the only significant structural difference between human CypA and CYP-3 being in the conformation adopted by Arg^{62} that places the side chain atoms 0.7Å apart.

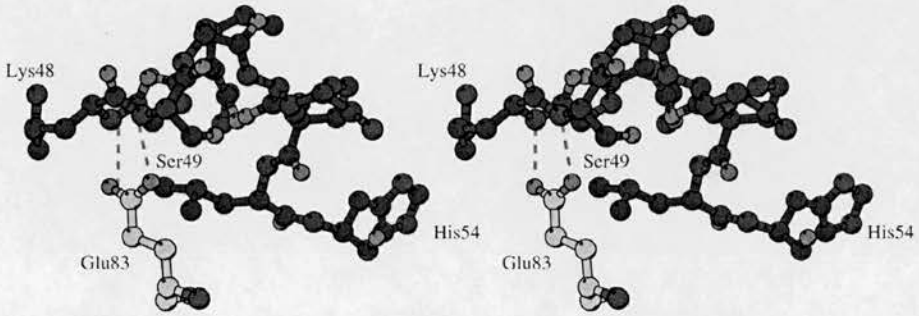


FIG. 4. **The divergent loop of *C. elegans* CYP-3.** The two strong hydrogen bonds from the carboxylic group of Glu^{83} to the amide main chain nitrogen atoms of Lys^{48} and Ser^{49} are shown as *violet*. All cyclophilin homologues with additional divergent loops have a conserved glutamate.

from the filarial nematode *Brugia malayi* (Fig. 1). The loop is held in place by a very specific network of hydrogen bonds from the side chain of Glu^{83}. The carboxylic group of Glu^{83} forms two strong hydrogen bonds with amide main chain nitrogen atoms of Lys^{48} and Ser^{49} (Fig. 4). All cyclophilin homologues with additional loops (including human/bovine cyclophilin 40 and some parasite and plant cyclophilins) have a conserved glutamate in a homologous position. Thus the role of Glu^{83} in these divergent loop structures is to both lock and tether the loop into a particular conformation. The loop provides a distinctive recognition feature of this subfamily of cyclophilins, which may prove to be important in protein-protein interactions.

Loop structures do however show some divergence as can be seen on comparing the structures from CYP-3 and *B. malayi*. The presence of inserts of 6–8 residues has been described previously in a number of members of this protein family, including many plant species (35) and is a particularly common feature of the nematode cyclophilins, being found in 7 of the 11 characterized cyclophilins from *C. elegans* (18). A consensus sequence for this additional loop based on multiple sequence alignments is (**GK*LH).

It is interesting to note that this loop also contains part of a conserved ATP/GTP-binding site motif A (A/G-*X*4-G-K-S/T) (36). The high degree of conservation of this inserted sequence over diverse species suggests the loop structure and its locking mechanism are an important functional part of the protein whose role remains to be elucidated, rather than a neutral addition.

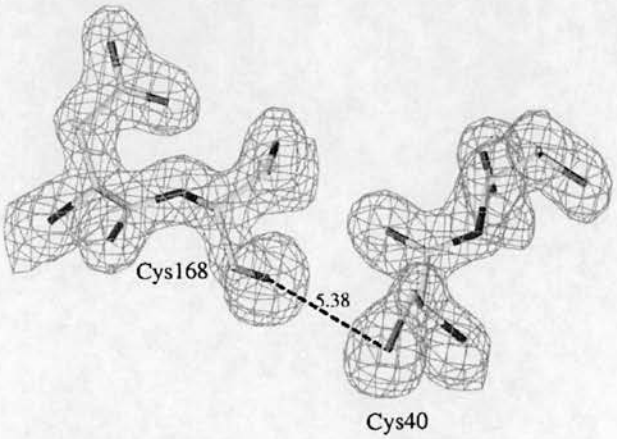


FIG. 5. **A 2 F_o-F_c electron density map around Cys^{40} and Cys^{168} is contoured at the 1.8σ level**. The distance between the two sulfur atoms is 5.38 Å. The well defined electron density clearly shows that both cysteine residues are present in the reduced form.

Conserved Cysteines—CYP-3 shares with all divergent loop cyclophilins, including Bm-CYP-1, two highly conserved cysteines (Fig. 1). The sulfur atoms of Cys^{40} and Cys^{168} are well defined in the electron density and are separated by 5.38 Å and are present in the structure as reduced cysteine residues (Fig. 5). Torsion angles of the 2 residues are N-Cα-Cβ-S = 53° for

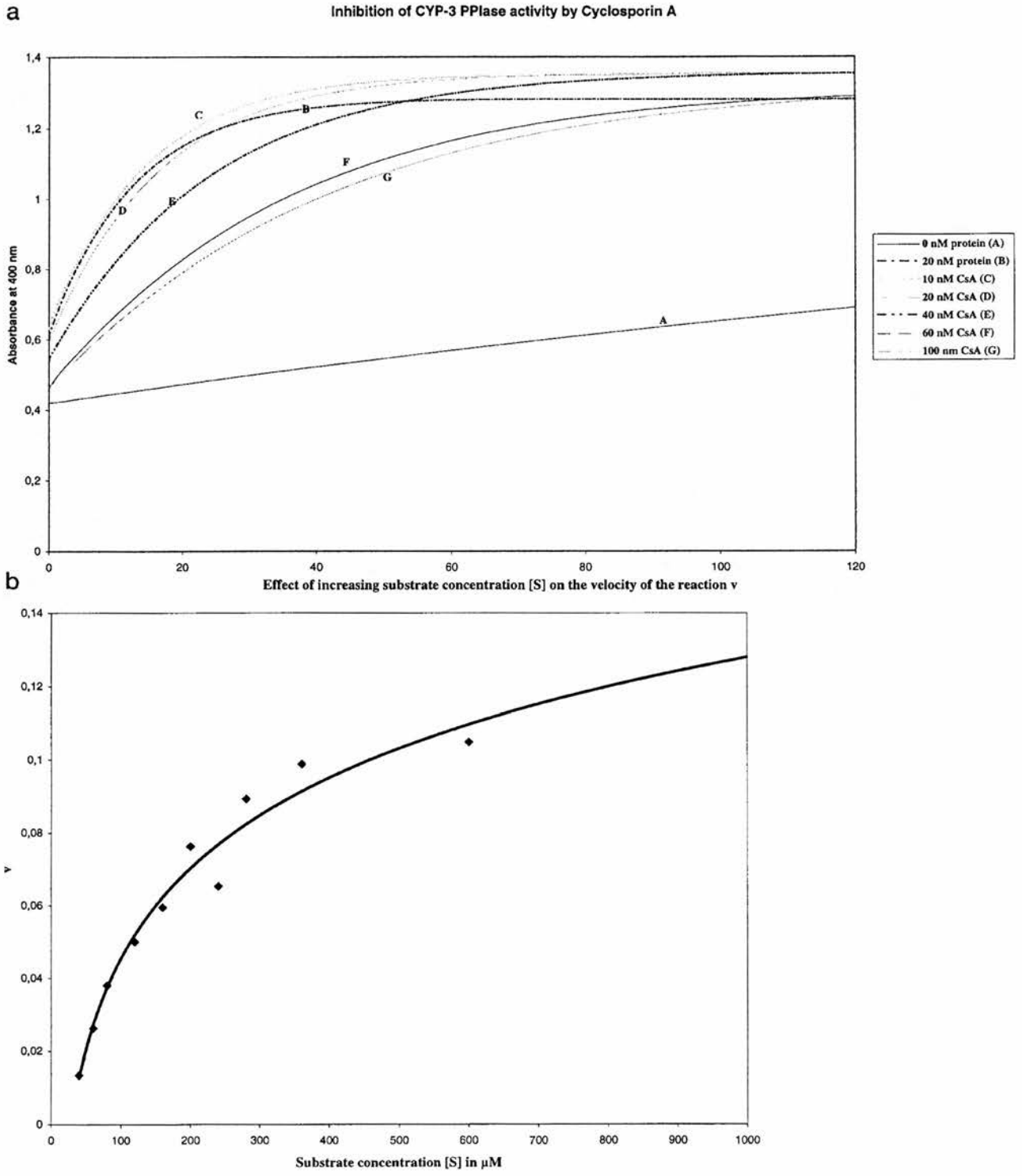


FIG. 6. *Panel a*, representative spectral progress curves showing ceCYP3 catalyzed acceleration of the rate of isomerization of the tetrapeptide substrate. The uncatalyzed rate (no CYP3) is labeled *A*. The curve labeled *B* shows the catalytic effect of 20 nM ceCYP3. Inhibition of the isomerization reaction by increasing concentrations of the cyclophilin binding drug cyclosporin A is illustrated by curves *C* through *G*. The spectral curves were obtained following preincubation with 10, 20, 40, 60, and 100 nM CsA are shown. Inhibition of PPIase activity can clearly be seen to increase with the amount of CsA added. Time is shown in seconds on *x*-axis. *Panel b*, the effect of increasing tetrapeptide substrate concentration on the velocity of the reaction. The ceCYP3 concentration was held constant at 20 nM, and substrate concentration was varied from 50 to 1,000 μM. It was not possible to measure rates with substrate concentration above 700 μM.

Cys^{168} and $-72°$ for Cys^{40}. The sulfur atom of Cys^{168} is 3.74Å away from nitrogen Ne2 of His^{54}, 3.2 Å from the carbonyl oxygen of Cys^{40} and 3.5 Å away from the backbone NH of Lys^{56}. Molecular modeling has shown that a simple rotation about the Cα-Cβ cysteine side chain bonds enables a disulfide bond to form. (N-Cα-Cβ-S = $-167°$ for Cys^{168} and N-Cα-Cβ-S = $-134°$

for Cys^{40}). These changes from the original staggered conformation to an alternatively favored staggered conformation results in an S-S distance of 2.1 Å, with no other short nonbonded interatomic contacts. The role of these conserved residues is unknown. The unexpected presence of the reduced form of these cysteine residues suggests the intriguing possibility that this feature may provide a signaling mechanism in response to oxidative stress. The formation of an S-S bond in an oxidizing environment may provide a dynamic signal inducing conformational change(s). Alternatively, the sulfhydryl group of the reduced Cys^{168} and Cys^{40} along with the imidazole side chains of His^{54} may provide a metal ion coordination site. Interestingly, His^{54} (or equivalent) is conserved in all these divergent loop cyclophilins (Fig. 1) and may also be an important part of this putative redox signaling mechanism.

Peptidyl-Prolyl cis-trans Isomerase Activity and Its Inhibition Activity by Cyclosporin A

Representative spectral progress curves showing the formation of the *trans*-isomer of the model substrate tetrapeptide can be seen in Fig. 6*a*. CYP-3 clearly accelerates the rate of isomerization of the tetrapetide substrate relative to the uncatalyzed thermal isomerization rate and the catalysis is inhibited by addition of the cyclophilin binding drug cyclosporin A. The enzyme reaction was found to follow Michaelis-Menten kinetics with the velocity of the reaction (v) increasing with substrate concentration [S] as shown in Fig. 6*b*. First order rate kinetics were observed and a double reciprocal Lineweaver-Burke plot of $1/v$ against $1/[S]$ gave values of $k_{cat} = 860\ s^{-1}$, $K_m = 350\ \mu M$, which corresponds to a value of k_{cat}/K_m of $2.1 \times 10^6\ M^{-1}\ s^{-1}$. These values are similar to the published values for human CypA, which has $k_{cat} = 12{,}700\ s^{-1}$, K_m 870 μM and $k_{cat}/K_m = 14.6 \times 10^6\ M^{-1}\ s^{-1}$ (24).

The spectral curves produced from the *in vitro* chymotrypsin-coupled peptide assay clearly shows CsA inhibition of the rate of CYP-3 catalyzed isomerization (Fig. 6*a*). A concentration of 16 nM CsA was required to inhibit 50% of the CYP-3 PPIase activity (IC_{50}). This value is comparable with the IC_{50} value of 19 nM obtained for the inhibition of human CypA by CsA (37).

Temporal and Spatial Expression Pattern of CYP 3

CYP 3 represents one of the most abundantly expressed cyclophilins in the nematode *C. elegans*. This observation is based both on the analysis of relative transcript abundance and due to the fact that it has been isolated 18 independent times during the *C. elegans* expression sequence tag project (38). The semi-quantitative RT-PCR analysis of the post-embryonic temporal expression pattern of *cyp-3* indicates that this transcript is most abundant during early larval development, peaking at the second larval stage, and dropping off in later development (Fig. 7). The spatial expression pattern of CYP-3 was examined by constructing a translational fusion of the promoter region with a reporter gene and staining the transgenic nematodes for β-galactosidase activity. This analysis revealed a very specific tissue localization for this highly conserved cyclophilin, being expressed exclusively in larval stages in the single celled anterior excretory system (Fig. 8, *A* and *B*). This reporter construct is tagged to a nuclear localization signal, hence the specific signal is observed in largest mononucleate cell of *C. elegans*, the excretory cell (39).

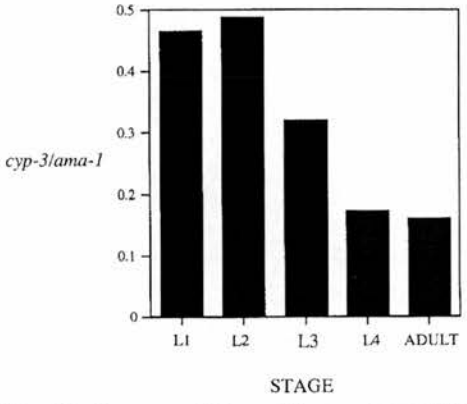


FIG. 7. **Graphical representation of the semi-quantitative RT-PCR result, illustrating the peak seen for the *cyp-3* transcript in the second larval stage**. The temporal abundance of *cyp-3* is expressed as a ratio to the constitutively expressed gene *ama-1*.

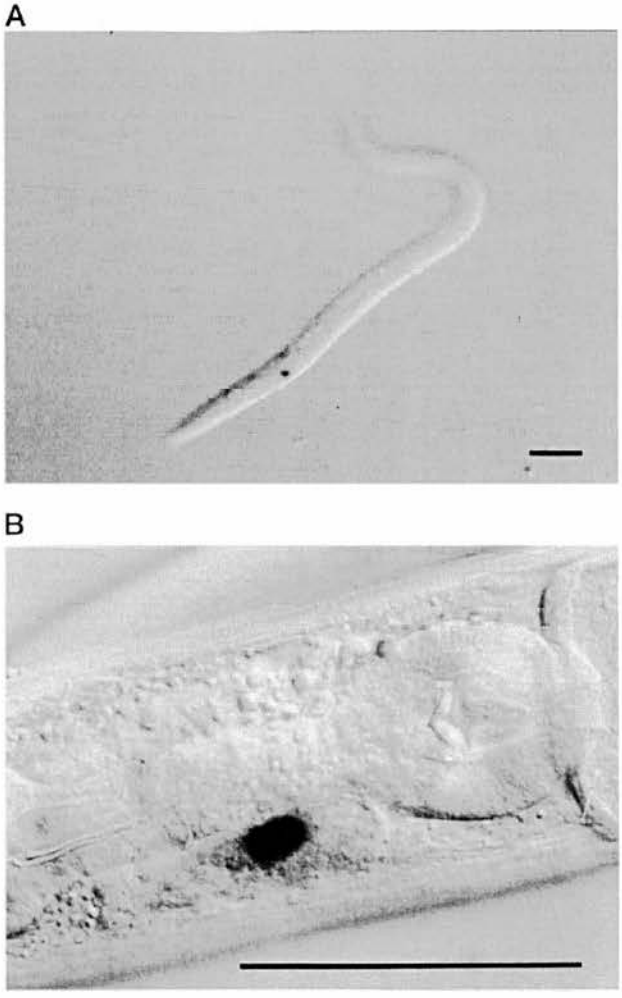


FIG. 8. **Micrograph showing the exclusive *cyp-3* reporter gene expression in *C. elegans***. Nuclear localized β-galactosidase activity is restricted to the single anterior mononuclear excretory cell. *A* and *B* represent individual larval nematodes. Bar represents 10 μM.

DISCUSSION

The x-ray structures of ceCYP3 together with sequence comparisons of the cyclophilin domains allow us to clearly define a subclass of divergent-loop cyclophilins, which have the following characteristics (highlighted in Fig. 1): an inserted loop (^{48}KSGKPLH^{54}); an invariant glutamate (Glu^{83}) and two unoxidized cysteines in close proximity (Cys^{40} and Cys^{168}). FASTA and BLAST searches of the available protein sequences (40, 41) using the consensus addition loop sequence GIGKSGKPLHFKG identified 50 related sequences, of which only 4 were

not divergent-loop cyclophilins. In each case, those cyclophilins that contained the loop sequence were found to contain two cysteines corresponding to Cys^{40} and Cys^{168} of CYP-3 and also Glu^{83}. Conversely, cyclophilins that do not have the additional loop have only one of the cysteine residues and also lack Glu^{83}. Our survey of cyclophilin sequences shows that plant cyclophilins in particular commonly contain the additional loop and concomitant pair of cysteine residues and some of these proteins have been linked to stress responses. For example *Solanum commersonii* significantly increases the production of sc-CyP under several forms of abiotic stress including cold shock, salicylic acid, and abscisic acid (42).

There may be a general link between cyclophilins and a cellular stress response as suggested by the association of human CypA with the thiol-specific antioxidant protein Aop1 (43). This activity is stimulated by the interaction between Aop1 with CypA and is not inhibited by CsA. One mechanism for the control of oxidative stress in bacteria is by the protein OxyR, which is activated by the formation of an intramolecular disulfide bond (44). A similar mechanism has been proposed for the activation of the Hsp33 chaperone protein, which is also redox regulated (45). Oxidation of Hsp33 by hydrogen peroxide causes disulfide bond formation, release of bound zinc, and activation of the chaperone activity of Hsp33.

The x-ray structures of CYP-3 described in this paper and the structure of CYP-1 from *B. malayi* (33, 34), another member of the divergent-loop subclass of cyclophilins, both show the invariant conserved cysteine residues in the reduced form. Furthermore, His^{54}, which is incorporated in the divergent loop, is located 3.6 Å from the sulfhydryl group of Cys^{168}. The arrangement of two cysteines and a histidine provides a suitable coordination for zinc. A very similar coordination geometry involving two cysteines and a histidine is present in the x-ray structure of a zinc complex of alcohol dehydrogenase (46). We suggest that the arrangement of cysteines and histidines, uniquely and invariantly present in this newly defined divergent-loop subclass, are likely to be involved in either metal coordination or in a redox-signaling process triggered by the formation of a Cys^{40} to Cys^{168} disulfide bond.

The entire genome of *C. elegans* has now been sequenced (38), and sequence searches show that there are a total of 14 *C. elegans* proteins that contain a cyclophilin-like domain. Three of these (CYP-3, CYP-2, and CYP-7) are single domain proteins with high sequence similarity to human cyclophilin A. It has been hypothesized (18) that the large number of CYP isoforms seen in *C. elegans* might be reflective of regulatory flexibility, providing interchangeable functions in different cell types, cellular compartments, or organelles. The expression pattern seen for the *cyp-3* transcript may correlate with a peak in the amount of newly synthesized structural proteins produced by the developing larvae. The exclusive excretory cell expression pattern is also intriguing. The excretory system in *C. elegans* is involved in osmoregulation and secretion (39). The role in osmoregulation is supported by gland cell laser ablation experiments, which result in worms becoming bloated and dying in hypotonic medium (47). The importance of excretory system in secretion is well established for parasitic nematodes, where it is the source of surface coat antigens for parasites such as *Toxocara canis* (48, 49). A role for CYP-3 in osmoregulation or secretion, however, remains to be established.

These expression pattern analyses coupled with the structural observations defining the new divergent-loop subgroup of cyclophilins, suggest that these cyclophilins may be involved in cellular responses to stress caused by changes in the redox environment or by periods of high cellular activity.

REFERENCES

1. Galat, A., and Metcalfe, S. M. (1995) *Prog. Biophys. Mol. Biol.* **63,** 67–118
2. Handschumacher, R. E., Harding, M. W., Rice, J., and Drugge, R. J. (1984) *Science* **226,** 544–547
3. Harding, M. W., Handschumacher, R. E., and Speicher, D. W. (1986) *J. Biol. Chem.* **261,** 8547–8555
4. Fischer, G., Wittmannliebold, B., Lang, K., Kiefhaber, T., and Schmid, F. X. (1989) *Nature* **337,** 476–478
5. Lang, K., Schmid, F. X., and Fischer, G. (1987) *Nature* **329,** 268–270
6. Takahashi, N., Hayano, T., and Suzuki, M. (1989) *Nature* **337,** 473–475
7. Bachinger, H. P. (1987) *J. Biol. Chem.* **262,** 17144–17148
8. Steinmann, B., Bruckner, P., and Superti-Furga, A. (1991) *J. Biol. Chem.* **266,** 1299–1303
9. Freskgard, P. O., Bergenhem, N., Jonsson, B. H., Svensson, M., and Carlsson, U. (1992) *Science* **258,** 466–468
10. Stamnes, M. A., Shieh, B. H., Chuman, L., Harris, G. L., and Zuker, C. S. (1991) *Cell* **65,** 219–227
11. Baker, E. K., Colley, N. J., and Zuker, C. S. (1994) *EMBO J.* **13,** 4886–4895
12. Ondek, B., Hardy, R. W., Baker, E. K., Stamnes, M. A., Shieh, B. H., and Zuker, C. S. (1992) *J. Biol. Chem.* **267,** 16460–16466
13. Helekar, S. A., and Patrick, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 5432–5437
14. Gothel, S. F., Scholz, C., Schmid, F., and Marahiel, M. A. (1998) *Biochemistry* **37,** 13392–13399
15. Rassow, J., Mohrs, K., Koidl, S., Barthelmess, I. B., Pfanner, N., and Tropschug, M. (1995) *Mol. Cell. Biol.* **15,** 2654–2662
16. Leverson, J. D., and Ness, S. A. (1998) *Mol. Cell* **1,** 203–211
17. Duina, A. A., Kalton, H. M., and Gaber, R. F. (1998) *J. Biol. Chem.* **273,** 18974–18978
18. Page, A. P., Macniven, K., and Hengartner, M. O. (1996) *Biochem. J.* **317,** 179–185
19. Chappell, L. H., and Wastling, J. M. (1992) *Parasitology* **105,** S 25-S 40
20. Page, A. P., Kumar, S., and Carlow, C. K. S. (1995) *Parasitol. Today* **11,** 385–388
21. Page, A. P., Landry, D., Wilson, G. G., and Carlow, C. K. S. (1995) *Biochemistry* **34,** 11545–11550
22. Page, A. P., and Winter, A. D. (1998) *Mol. Biochem. Parasitol.* **95,** 215–227
23. Johnstone, I. L., Shafi, Y., Majeed, A., and Barry, J. D. (1996) *Mol. Biochem. Parasitol.* **80,** 103–112
24. Kofron, J. L., Kuzmic, P., Kishore, V., Colonbonilla, E., and Rich, D. H. (1991) *Biochemistry* **30,** 6127–6134
25. Otwinowski, Z., and Minor, W. (1993) *Methods Enzymol.* **276,** 307–326
26. Navaza, J. (1994) *Acta Crystallogr.* **A 50,** 157–163
27. Sheldrick, G. M. (1997) *SHELX97* and *SHELXWAT*, University of Göttingen, Germany
28. Braun, W., Kallen, J., Mikol, V., Walkinshaw, M. D., and Wuthrich, K. (1995) *FASEB J.* **9,** 63–72
29. Taylor, P., Husi, H., Kontopidis, G., and Walkinshaw, M. D. (1997) *Prog. Biophys. Mol. Biol.* **67,** 155–181
30. Kallen, J., Spitzfaden, C., Zurini, M. G. M., Wider, G., Widmer, H., Wuthrich, K., and Walkinshaw, M. D. (1991) *Nature* **353,** 276–279
31. Ke, H. M., Zydowsky, L. D., Liu, J., and Walsh, C. T. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9483–9487
32. Pflugl, G. M., Kallen, J., Jansonius, J. N., and Walkinshaw, M. D. (1994) *J. Mol. Biol.* **244,** 385–409
33. Taylor, P., Page, A. P., Kontopidis, G., Husi, H., and Walkinshaw, M. D. (1998) *FEBS Lett.* **425,** 361–366
34. Mikol, V., Ma, D., and Carlow, C. K. S. (1998) *Prot. Sci.* **7,** 1310–1316
35. Chou, I. T., and Glasser, C. S. (1997) *Plant Mol. Biol.* **35,** 873–892
36. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J.* **1,** 945–951
37. Liu, J., and Walsh, C. T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 4028–4032
38. Anonymous (1999) *Science* **283,** 35
39. White, J. (1988) in *The Nematode Caenorhabditis elegans* (Wood, W. B. E., ed) pp. 81–123, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
40. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 2444–2448
41. Altschul, S. F., and Gish, W. (1996) *Methods Enzymol.* **266,** 460–480
42. MezaZepeda, L. A., Baudo, M. M., Palva, E. T., and Heino, P. (1998) *J. Exp. Bot.* **49,** 1451–1452
43. Jaschke, A., Mi, H. F., and Tropschug, M. (1998) *J. Mol. Biol.* **277,** 763–769
44. Zheng, M., Aslund, F., and Storz, G. (1998) *Science* **279,** 1718–1721
45. Jakob, U., Muse, W., Eser, M., and Bardwell, J. C. A. (1999) *Cell* **96,** 341–352
46. Cho, H., Ramaswamy, S., and Plapp, B. V. (1997) *Biochemistry* **36,** 382–389
47. Nelson, F. K., and Riddle, D. L. (1984) *J. Exp. Zool.* **231,** 45–46
48. Page, A. P., Rudin, W., Fluri, E., Blaxter, M. L., and Maizels, R. M. (1992) *Exp. Parasitol.* **75,** 72–86
49. Page, A. P., Hamilton, A. J., and Maizels, R. M. (1992) *Exp. Parasitol.* **75,** 56–71
50. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
51. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950

Experimental Section

A series of Mo/HZSM-5 catalysts were prepared as described in ref. [4b]. HZSM-5 powders were impregnated with aqueous solutions containing a given amount of ammonia heptamolybdate (AHM), then dried at RT and 373 K for 12 and 8 h, respectively. After calcined at 773 K for 5 h, the catalysts were crushed and sieved to yield granules of 20–60 mesh.

A specially designed device was built for the MAS NMR study. It contains a gas flow section, a reaction and sample preparation section, and an on-line mass spectrometer (Figure 1). After the reaction (the temperature can be as high as 1123 K), the reaction and sample preparation section can be taken off and turned through 90°. Then, in the reaction atmosphere the sample can be transferred into a NMR-rotor with the aid of a tamper (to pack the sample into the tube) that is mounted in a bellows attatched to the upper side of the flange. The rotor is then sealed with a cap in a rack with the aid of tamper. A spin rate as high as 12 kHz can be reached following this sample filling method. In the present case, the samples were treated with a methane flow ($SV = 1500\ mL\,g^{-1}\,h^{-1}$, 1 atm) for 1 h at 573, 673, and 873 K. Subsequently, the catalysts were subjected to reactions of 10 min, 0.5 h, 1 h, 3 h, and 6 h. In all the cases, the temperature was raised at a rate of $15\,^{\circ}C\,min^{-1}$.

Proton MAS NMR spectra were recorded in 4 mm ZrO_2 rotors at 400.13 MHz on a Bruker DRX-400 spectrometer fitted with a MAS probe. For each spectrum, a $\pi/10$ pulse and a 4 s relaxation delay were used, and 400 scans were accumulated. Samples were spun at 8 kHz, and the DSS (DSS = sodium 4,4-dimethyl-4-silapentane sulfonate) signal was taken as the reference for the ^1H chemical shifts. The deconvolution of the spectra was conducted using WINNMR, supplied by the spectrometer manufacturer.

Received: March 9, 2000 [Z14824]

[1] a) Y. Ono, *Catal. Rev. Sci. Eng.* **1992**, *34*, 179; b) P. Meriaudeau, C. Naccache, *Catal. Rev. Sci. Eng.* **1997**, *39*, 5.

[2] G. Buckles, G. J. Hutchings, C. D. Williams, *Catal. Lett.* **1991**, *11*, 89.

[3] L. Wang, L. Tan, M. Xie, G. Xu, J. Huang, Y. Xu, *Catal. Lett.* **1993**, *21*, 35.

[4] a) Y. Xu, S. Liu, L. Wang, M. Xie, X. Guo, *Catal. Lett.* **1995**, *30*, 135; W. Liu, Y. Xu, *J. Catal.* **1999**, *185*, 386; b) Y. Shu, Y. Xu, S.-T. Wong, L. Wang, X. Guo, *J. Catal.* **1997**, *170*, 11; c) D. Wang, J. H. Lunsford, M. P. Rosynek, *J. Catal.* **1997**, *169*, 347; d) B. M. Weckhuygsen, D. Wang, M. P. Rosynek, J. H. Lunsford, *Angew. Chem.* **1997**, *109*, 2471; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2374; e) S. Liu, L. Wang, R. Ohnishi, M. Ichikawa, *J. Catal.* **1999**, *181*, 175; f) F. Solymosi, J. Cserenyi, A. Szoke, T. Bansagi, A.Oszko, *J. Catal.* **1997**, *165*, 150; g) F. Solymosi, L. Bugyi, A. Oszko, *Catal. Lett.* **1999**, *57*, 103.

[5] H. Ernst, D. Freude, T. Mildner, I. Wolf, *Stud. Surf. Sci. Catal.* **1995**, *94*, 413; T. Mildner, D. Freude, *J. Catal.* **1998**, *178*, 309.

[6] M. Hunger, T. Horvath, *J. Chem. Soc. Chem. Commum.* **1995**, 1423; M. Hunger, M. Seiler, T. Horvath, *Catal. Lett.* **1999**, *57*, 199; P. K. Isbester, A. Zalusky, D. H. Lewis, M. C. Douskey, M. J. Pomije, K. R. Mann, J. Munson, *Catal. Today* **1999**, *49*, 363.

[7] E. MacNamara, D. Raftery, *J. Catal.* **1998**, *175*, 135.

[8] J. F. Haw, B. R. Richardson, I. S. Oshiro, N. D. Lazo, J. S. Speed, *J. Am. Chem. Soc.* **1989**, *111*, 2052; J. Munson, J. F. Haw, *J. Am. Chem. Soc.* **1991**, *113*, 6303; J. Munson, D. Murray, J. F. Haw, *J. Catal.* **1993**, *141*, 733; J. F. Haw, T. Xu, J. B. Nicholas, P. W. Goguen, *Nature* **1997**, *389*, 23; J. F. Haw, P. W. Goguen, T. Xu, T. W. Skloss, W. Song, Z. Wang, *Angew. Chem.* **1998**, *110*, 1001–1005; *Angew. Chem. Int. Ed.* **1998**, *37*, 948.

[9] M. Hunger, *Catal. Rev. Sci. Eng.* **1997**, *39*, 345.

[10] M. Hunger, D. Freude, H. Pfeifer, *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 657.

[11] V. R. Choudhary, A. K. Kinage, T. V. Choudhary, *Science* **1997**, *275*, 1286.

[12] In the TPSR (temperature program surface reaction) profile of this reaction (detected by the on-line mass spectrometer) water was shown to be produced when the temperature reached 973 K, with a maximum at about 13 min after the reaction began. Two minutes later (15 min after the start of the reaction) the start of benzene formation is observed. The formation of water stopped about 1 h later at the same temperature, but that of benzene increases slowly in the first hour, and then appears to be constant. This indicates that the water formation is a stoichiometric (not catalytic) reaction [see Eq. (1)], where as benzene formation is a catalytic reaction.

[13] C. C. Liu, G. E. Maciel, *Anal. Chem.* **1996**, *68*, 1401.

First Crystal Structure of a Medicinally Relevant Gold Protein Complex: Unexpected Binding of $[Au(PEt_3)]^+$ to Histidine**

Juan Zou, Paul Taylor, Jacqueline Dornan, Stephen P. Robinson, Malcolm D. Walkinshaw,* and Peter J. Sadler*

Antiarthritic gold(I) compounds such as the injectable drugs aurothiomalate[1] (Myochrysine) and aurothioglucose (Solganol),[2] and the oral drug[3] $[Au^I(PEt_3)$ (2,3,4,6-tetra-*O*-acetyl-β-D-thioglucose-*S*)] (auranofin) are "prodrugs", which undergo facile ligand displacement reactions.[4, 5] The critical target sites are thought to be thiolate sulfurs (cysteine residues) in proteins and enzymes,[5] Au^I being a very "soft" (class 'b') metal ion.[6, 7] However, there is a lack of structural data on adducts of antiarthritic Au^I complexes with proteins.[8] We report here investigations of reactions between the enzyme cyclophilin and the antiarthritic complex[9, 10] $[Au(PEt_3)Cl]$ (**1**), and the first X-ray crystal structure of a protein adduct with a Au^I phosphane complex. Unexpectedly, in cyclophilin crystals, Au^I binds to the nitrogen atom of an active site His residue, despite the presence of four Cys thiol groups. The results have implications for understanding the biological chemistry of gold antiarthritic complexes. Additional interest in this work arises from the need for site-specific heavy-atom reagents for X-ray crystallography.[8]

[*] Prof. Dr. M. D. Walkinshaw, Dr. P. Taylor, J. Dornan, S. P. Robinson
Institute of Cell and Molecular Biology
Michael Swann Building
The University of Edinburgh, King's Buildings
West Mains Road, Edinburgh EH9 3JR (UK)
Fax: (+44) 131-650-7055
E-mail: M.Walkinshaw@ed.ac.uk

Prof. Dr. P. J. Sadler, Dr. J. Zou
Department of Chemistry
The University of Edinburgh, King's Buildings
West Mains Road, Edinburgh EH9 3JJ (UK)
Fax: (+44) 131-650-6452
E-mail: P.J.Sadler@ed.ac.uk

[**] We thank the Wellcome Trust (Fellowship for J.Z.), BBSRC, CLRC, and EPSRC for their support for this work, Dr. Adam Gouldsworthy and Mr. Mark Scott (Edinburgh Centre for Protein Technology) for their assistance with LC-ESI-MS, Drs. George Kontopidis, Su-ying Wu, and Hongzhe Sun for their valuable suggestions, discussion, and technical assistance.

Supporting information for this article is available on the WWW under http://www.wiley-vch.de/home/angewandte/ or from the author.

  1433-7851/00/3916-2931 $ 17.50+.50/0

Cyclophilins are ubiquitous proteins involved in protein folding, transport, and assembly, and are linked to general cellular stress responses.[11] They are peptidyl prolyl isomerase (PPIase) enzymes, and cyclophilin A (an immunophilin) is the major intracellular receptor for the immunosuppressive drug, cyclosporin A (CsA). We chose cyclophilin-3 (Cyp-3, one of the most abundantly expressed cyclophilin isoforms in the free living nematode *Caenorhabditis elegans*) for study, since it has a known X-ray structure[12] and possesses four free Cys thiol groups which are potential Au^I target sites.

First the reaction between Cyp-3 and **1** was studied in solution by mass spectrometry. For a molar ratio of Cyp-3:**1** of 1:3.5, the sample after one day revealed three new peaks with masses of 18612.0 u (calcd 18615.5 u), 18736.8 u (calcd 18733.6 u), and 18932.9 u (calcd 18930.6 u), corresponding to {Cyp-3 + Au}, {Cyp-3 + [$AuPEt_3$]}, and {Cyp-3 + Au + [$AuPEt_3$]}, respectively (Figure 1). For a molar ratio Cyp-3:**1** of 1:7, the sample after five days revealed that the peaks for

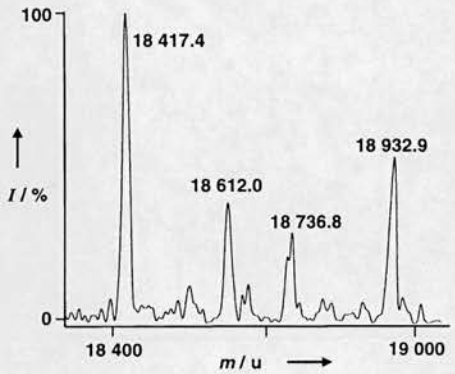


Figure 1. LC-ESI-MS of Cyp-3 (13.5 μM) after incubation with **1** (47.3 μM) for 24 h at 291 K. The peaks correspond to Cyp-3 (18417.4 u, theoretical mass: 18419.2 u); {Cyp-3 + Au} (18612.0 u, theoretical mass: 18615.5 u); {Cyp-3 + [$Au(PEt_3)$]} (18736.8 u, theoretical mass: 18733.6 u); {Cyp-3 + Au + [$Au(PEt_3)$]} (18932.9 u, theoretical mass: 18930.6 u).

Cyp-3 and {Cyp-3 + [$AuPEt_3$]} had disappeared, peaks for {Cyp-3 + Au} (18618.8 u), and {Cyp-3 + Au + [$AuPEt_3$]} (18930.8 u, major peak) were still detectable, and new peaks appeared at 19242.5 u corresponding to {Cyp-3 + Au + 2[$AuPEt_3$]} (calcd 19245.7 u) and 18812.8 u corresponding to {Cyp-3 + 2Au} (calcd 18812.4 u). These results appeared to be consistent with chloride displacement and $[AuPEt_3]^+$ binding to the protein, followed by PEt_3 displacement (and oxidation to $OPEt_3$). Such a displacement might be facile in the region of Cys40 and Cys168 (highly conserved in all divergent loop cyclophilins) since their sulfur atoms are within 5.5 Å and could readily provide linear S-Au-S coordination for Au^I.

Reactions of Cyp-3 with the disulfide 5,5′-dithio-bis(2-nitrobenzoic acid (DTNB) also suggested that all four Cys thiol groups are accessible. By UV/Vis spectroscopy, the average thiol content of Cyp-3 was determined to be 4.2 mol thiol per mol Cyp-3. The reaction was also investigated by LC-ESI-MS. The chromatographic peak for Cyp-3 alone corresponded to a mass of 18418.5 u (mean value, theoretical mass 18419.2 u).[13] This peak disappeared after reaction of Cyp-3 with DTNB, and the major new chromatographic peak had a mass of 19210.6 u (Figure 2), corresponding to {Cyp-3 + 2DTNB} (i.e. 4 × 1/2 DTNB, calcd: 19211.2 u). This implies that all four Cys thiol groups form mixed disulfides with DTNB and are therefore accessible.

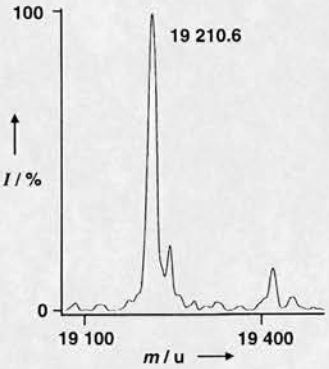


Figure 2. LC-ESI-MS of Cyp-3 after reaction with excess DTNB. The major peak (19210.6 u) corresponds to four half DTNB species bound to Cyp-3, that is all four Cys residues blocked as mixed disulfides (theoretical mass 19211.2 u).

To identify the gold coordination sites, we treated single crystals of Cyp-3 with complex **1** for periods of 0.42, 2, 4, 6 h, and two weeks. The crystals were then frozen and examined by X-ray diffraction. All crystals soaked with **1** changed their appearance. They became cracked and lost their sharply defined edges. A crystal structure determined after 0.42 h soaking did not show the presence of gold in the calculated electron density map, and after soaking for 6 h did not diffract at all. After samples had been soaked for two weeks, cracking was severe. The electron density map from crystals soaked for 2 or 4 h contained a large peak corresponding to a single gold site (Figure 3). The structure was solved by using isomorphous replacement since that of unliganded Cyp-3 is known.[12] In the first difference map, Au was clearly seen together with additional electron density in the region expected for a ligand *trans* to His133 Nε2. This was interpretable as PEt_3 as indicated by the shape of the electron density. The structure was refined with SHELXL97[14] using 20154 unique data (5% for the calculation of R_{free}) and initial constraints on the Au–N and Au–P bond lengths of 2.05 and 2.27 Å, respectively, values similar to those in crystalline [Au(2-isopropylimidazole)(Pcy_3)][15] and [Au(phthalimido)(PEt_3)][16]. The final R factor was 18.48% and R_{free} 24.68%.

Surprisingly, the gold ion was not bonded to a Cys sulfur atom, but instead to the Nε2 nitrogen atom of His133 (Au–N 2.06 Å) and to PEt_3 (Au–P 2.27 Å). The linear geometry (N-Au-P angle 167.7°) is characteristic of Au^I. The temperature factors for the complex (Nε2 13.86, Au 43.77, P 79.45, C(av) 74.96) suggest that the PEt_3 group is undergoing rotational motion.

Of the four His residues in Cyp-3, the imidazole rings of only His54 and His133 are solvent accessible, both in a solvent channel running between two Cyp-3 molecules related by a twofold screw axis parallel to the *b* axis. Cys168 also appears to be readily accessible and the sulfur atom of Cys163

  1433-7851/00/3916-2932 $ 17.50+.50/0

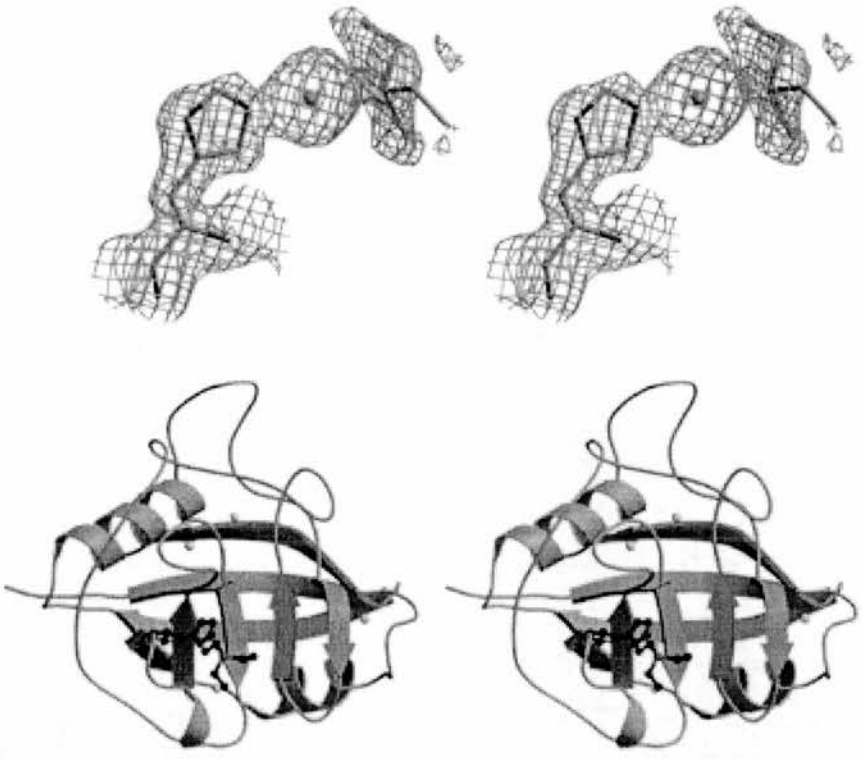


Figure 3. Top: Refined structure of the gold binding site in Cyp-3, showing His133 bound to PEt_3. The $2F_o - F_c$ electron density map is contoured at 1.75σ. The lack of definition of the PEt_3 ligand is probably due to rotational disorder. Bottom: Stereoview of Cyp-3 after reaction with **1** showing the S atoms of the four Cys residues (light spheres: Cys40 (top left), Cys168 (top right), Cys122 (bottom left), Cys163 (bottom right)) and $[AuPEt_3]^+$ bound to the Nε2 center of His133. Drawn using Bobscript,[26] MolScript,[27] and Raster3D.[28]

is visible from the protein surface, but these sites may not be able to accommodate $[AuPEt_3]^+$ without severe structural distortions and disruption of the crystal.

Gold binding at His133 should inhibit Cyp-3 since this is an active-site residue, and this was found to be the case. A chymotrypsin-coupled assay[17] of the isomerization of the tetrapeptide substrate Suc-Ala-Ala-Pro-Phe-*p*-nitroanilide (Bachem) showed that the concentration of **1** required to inhibit 50% of Cyp-3 PPIase activity was 14 nM (Figure 4), a value comparable with that for inhibition of human CypA by CsA.[18]

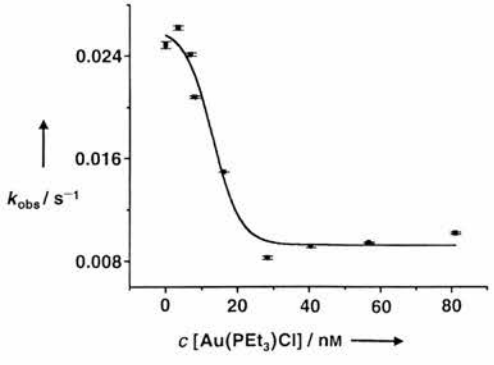


Figure 4. Variation of the observed rate constant for PPIase activity of Cyp-3 with increasing concentration of **1**. An IC_{50} value of 14 nM was calculated from this curve.

Au^I has a moderate affinity for N ligands in low molecular mass complexes,[19] but a much higher affinity for thiolate groups. Hence a Cys-S atom is much more effective than an imidazole-N atom in displacing the Cl^- ion from complex **1** in aqueous solution.[20] Complex **1** forms adducts with His residues of serum albumin only after the thiolate S atom at Cys34 has been saturated with Au^I, first as the mono-adduct Cys34-S-($AuPEt_3$), and then as a bisadduct Cys34-S-$(AuPEt_3)_2$.[2, 4, 21, 22] Such a bisadduct could also form with Cyp-3 in the presence of excess **1**.

This appears to be the first X-ray structural characterization of an adduct of a gold–phosphane complex with a His residue in a protein.[8] Previously it was suggested[23] that Au^I–His interactions are responsible for the ability of complex **1** to induce low-spin- to high-spin-state transitions of heme proteins such as cytochrome c, but direct structural evidence for this was not obtainable. The present results confirm that His imidazole rings can indeed be target sites for $[AuPEt_3]^+$ binding in proteins.

Experimental Section

[Au(PEt_3)Cl] (**1**) was purchased from Strem Chemicals and Alfa-Johnson Matthey plc. Cyclophilin-3 (Cyp-3) was expressed and purified as described previously.[12]

ESI-HPLC-MS: Cyp-3 in Tris buffer (10 mM; pH 7.3), or in HEPES buffer (20 mM; pH 7.0), was treated with complex **1** (3.5 or 7 molar equivalents of a 2.5 mM stock solution in ethanol or methanol, diluted to 150 mM with buffer before use), and incubated for 1–5 days at 291 K. Prior to ESI-MS, the protein was desalted, separated from free gold complex on a Phenomenex (Jupiter) C_{18} reverse-phase column, and eluted with H_2O/acetonitrile/TFA (0.05% v/v). Positive-ion electrospray mass spectrometry was performed on a Platform II instrument (Micromass), with source temperature of 413 K and data acquisition in continuum mode, scan rate 900 u s^{-1}. The data were processed with a Mass Lynx (version 2.3) Windows NT PC data system using a maximum entropy software algorithm. Mass accuracy was within 0.1 *m/z* unit.

Determination of free SH groups: Cyp-3 (14.0 μM, concentration determined by Bradford assay) was incubated with DTNB (5 mM) in HEPES buffer (20 mM; pH 7.0) for 30 min at 291 K, and the absorption was measured at 412 nm.[24, 25]

Crystallization, Au^I soaking, and data collection: Tetragonal crystals of Cyp-3 were grown using the hanging-drop vapour diffusion method: Cyp-3 (16 mg mL^{-1}) in sodium citrate (50 mM; pH 5.6), 16–17.5% wt/v methoxypolyethylene glycol 5000 (MPEG 5000) (4 μL) suspended over sodium citrate (100 mM), 32–35% MPEG 5000 at 291 K. Average crystal dimensions: 0.5 × 0.4 × 0.4 mm. For crystal soaking, an aliquot of a stock solution of **1** in DMSO was diluted with sodium citrate (100 mM) containing 33–35% MPEG 5000, pH 5.6, giving a final Au^I concentration of 3.5 mM with 5% DMSO present. The stock solution of **1** in DMSO gave rise to only one

  1433-7851/00/3916-2933 $ 17.50+.50/0

$^{31}P\{^1H\}$ NMR peak at $\delta = 34.5$ when monitored over a period of one month, showing that the Au–P bond is stable under these conditions. Crystals were transferred from hanging drops with 0.4–0.5 mm cryoloops into cryoprotectant solution (3 μL, mother liquor plus 20% glycerol), and 5 s after removal from cryoprotectant were mounted in a cryoloop (Hampton Research) and flash frozen in liquid N_2. X-ray data were collected at 100 K (Oxford Cryosystems) using MAR image plates at the Daresbury synchrotron source, and processed using DENZO. The data set consists of 185907 measured reflections and provides a unique data set of 20154 reflections with an $R_{merge} = 0.068$ and an overall completeness to 1.85 Å of 0.983. R_{merge} in the 1.88–1.85 Å resolution shell is 0.253.

PPIase assay: Cyp-3 was incubated with **1** at 277 K at molar ratios of **1**: Cyp-3: of 0, 0.43, 0.86, 1.0, 2.0, 3.5, 5.0, 7.0, and 10.0, for 48 h in HEPES buffer (50 mM) containing NaCl (86 mM), pH 8.0, and 4 μL aliquots were then assayed at 275 K following the procedure previously described.[12]

ESI-MS spectra of Cyp-3 after reaction with Et_3PAuCl (**1**) for five days , a colour version of Figure 3, and plots of absorbance versus time for assay of the PPIase activity of Cyp-3 are available in the Supporting Information. The atomic coordinates (code 1E3B) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Received: February 21, 2000 [Z14741]

[1] R. Bau, *J. Am. Chem. Soc.* **1998**, *120*, 9380–9381.
[2] M. T. Coffer, C. F. Shaw III, M. K. Eidsness, J. W. Watkins II, R. C. Elder, *Inorg. Chem.* **1986**, *25*, 333–339.
[3] D. T. Hill, B. M. Sutton, *Cryst. Struct. Commun.* **1980**, *9*, 679–686.
[4] C. F. Shaw III, *Chem. Rev.* **1999**, *99*, 2589–2600.
[5] M. T. Razi, G. Otiko, P. J. Sadler, *Am. Chem. Soc. Symp. Ser.* **1983**, *209*, 371–384.
[6] R. G. Pearson, *J. Am. Chem. Soc.* **1963**, *85*, 3533–3539.
[7] S. Ahrland, J. Chatt, N. R. Davies, *Q. Rev. Chem. Soc.* **1958**, *12*, 265–276.
[8] The most comprehensive database of Au sites in protein crystals is : http://www.bmm.icnet.uk/had/heavyatom.html. Most are $[Au(CN)_2]^-$ sites.
[9] Complex **1** exhibits a) antiarthritic activity: B. M. Sutton, E. McGusty, D. T. Walz, M. J. DiMartino, *J. Med. Chem.* **1972**, *15*, 1095–1098; b) anticancer activity: C. K. Mirabelli, R. K. Johnson, D. T. Hill, L. F. Faucette, G. R. Girard, G. Y. Kuo, C. M. Sung, S. T. Crooke, *J. Med. Chem.* **1986**, *29*, 218–223.
[10] Complex **1** is a major product a) from reactions of the oral drug auranofin in HCl solutions: J. Hempel, Y. Mikuriya in *Bioinorganic Chemistry of Gold Coordination Compounds* (Eds.: B. M. Sutton, R. G. Franz), SK&F, Philadelphia, **1981**, pp. 37–46; b) from reactions of auranofin with hypochlorite, an oxidant present in inflamed tissues: A. J. Canumalla, PhD thesis, University of Wisconsin-Milwaukee (USA), **1998**; c) ref. [4].
[11] A. Galat, S. M. Metcalfe, *Prog. Biophys. Mol. Biol.* **1995**, *63*, 67–118.
[12] J. Dornan, A. P. Page, P. Taylor, S.-Y. Wu, A. D. Winter, H. Husi, M. D. Walkinshaw, *J. Biol. Chem.* **1999**, *274*, 34877–34883. Cyp-3 (173 amino acids) has a high sequence and structural similarity to human Cyp-A.
[13] The expressed protein lacks Met1.
[14] G. M. Sheldrick, **1997**, SHELXL-97, University of Göttingen, Germany.
[15] B. Bovio, F. Bonati, A. Burini, B. R. Pietroni, *Z. Naturforsch.* **1984**, *39B*, 1747–1754.
[16] S. J. Berners-Price, M. J. Dimartino, D. T. Hill, R. Kuroda, M. A. Mazid, P. J. Sadler, *Inorg. Chem.* **1985**, *24*, 3425–3434.
[17] J. L. Kofron, P. Kuzmic, V. Kishore, E. Colonbonilla, D. H. Rich, *Biochemistry* **1991**, *30*, 6127–6134.
[18] J. Liu, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4028–4032.
[19] J. Strähle in *Gold: Progress in Chemistry, Biochemistry and Technology* (Ed.: H. Schmidbaur), Wiley, Chichester, **1999**, pp. 311–348.
[20] a) M. T. Razi, PhD thesis, University of London, **1983**; b) M. T. Razi, P. J. Sadler, unpublished results. ^{35}Cl NMR studies show that one molar equivalent of L-Cys is sufficient to displace all the bound Cl^- ions from **1**, whereas more than eight molar equivalents of *N*-methylimidazole are required.
[21] N. A. Malik, G. Otiko, P. J. Sadler, *J. Inorg. Biochem.* **1980**, *12*, 317–322.
[22] C. F. Shaw III, N. A. Schaeffer, R. C. Elder, M. K. Eidsness, J. M. Trooster, G. H. M. Calis, *J. Am. Chem. Soc.* **1984**, *106*, 3511–3521.
[23] a) G. Otiko, P. J. Sadler, *FEBS Lett.* **1980**, *116*, 227–30; b) M. C. Grootveld, G. Otiko, P. J. Sadler, R. Cammack, *J. Inorg. Biochem.* **1986**, *27*, 1–15.
[24] A. F. S. A. Habeeb, *Methods Enzymol.* **1972**, *25*, 457–464.
[25] P. C. Jocelyn, *Methods Enzymol.* **1987**, *143*, 44–67.
[26] R. M. Esnouf, *J. Mol. Graphics Modelling* **1997**, *15*, 132–136.
[27] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946–950.
[28] E. A. Merritt, D. J. Bacon, *Methods Enzymol.* **1997**, *277*, 505–524.

DeNO*x* of Exhaust Gas from Lean-Burn Engines through Reversible Adsorption of N_2O_3 in Alkali Metal Cation Exchanged Faujasite-Type Zeolites**

Asima Sultana, Raf Loenders, Orietta Monticelli, Christine Kirschhock, Pierre A. Jacobs, and Johan A. Martens*

Nitrogen oxide (NO_x) adsorbents are key control components for the tailpipe emissions of transport vehicles powered with lean-burn engines. Current adsorbents are basic oxides which adsorb NO_x as nitrates, but suffer from poisoning by sulfur oxides. We discovered that NO and NO_2 can be selectively trapped as dinitrogen trioxide in alkali metal exchanged faujasite zeolites above 200 °C. The trapped N_2O_3 molecules compete with water molecules on a specific adsorption site and can be displaced by a change in partial pressure. Dinitrogen trioxide adsorption is not affected by the presence of sulfur oxides.

Modern lean-burn internal combustion engines operated with excess air show reduced fuel consumption and carbon dioxide emission but produce excessive amounts of nitrogen oxides (NO_x). The emission of NO_x at the tailpipe is the major source of pollution from transport vehicles powered by hydrocarbon fuels.[1] Since there is a trade-off between NO_x and particulate carbon formation in the lean-burn engine, an efficient NO_x-elimination system for the exhaust can offer a solution to the particulate carbon emissions at the same time.[2] In the emerging technologies that focus on reducing the NO_x emission from lean-burn engines, temporary storage

[*] Prof. Dr. ir. J. A. Martens, Dr. A. Sultana, Dr. R. Loenders, Dr. O. Monticelli, Dr. C. Kirschhock, Prof. Dr. ir. P. A. Jacobs
Centrum voor Oppervlaktechemie en Katalyse
Katholieke Universiteit Leuven
Kardinaal Mercierlaan 92, 3001 Heverlee (Belgium)
Fax: (+32) 16-321998
E-mail: johan.martens@agr.kuleuven.ac.be

[**] This work was sponsored by the European Community (Brite Euram III projects BE-95-2127 "SNR Technique" and BE-97-4493 "SORPTEC") and the Belgium government (IUAP-PAI program).

Supporting information for this article is available on the WWW under http://www.wiley-vch.de/home/angewandte/ or from the author.

  1433-7851/00/3916-2934 $ 17.50+.50/0

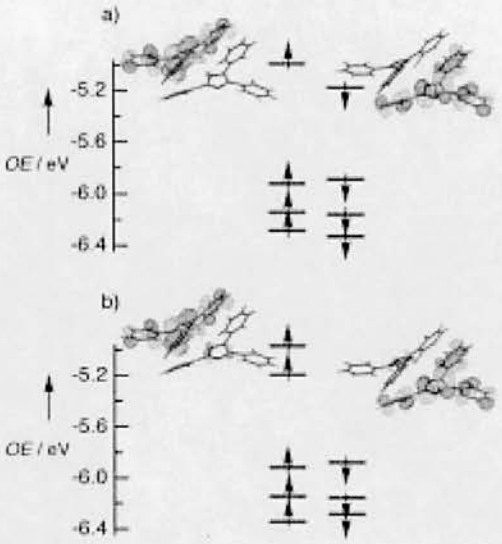


Figure 4. Molecular orbital energy level diagrams and singly occupied molecular orbitals (SOMOs) for a) the singlet state of pair B and b) the triplet state of pair B as obtained from the UB3LYP/6-31G* method. OE = orbital energy.

Table 1. Becke3LYP/6-31G* energies and $\langle S^2 \rangle$ values for the singlet and triplet states for pair B.

Spin state	HF energy (UB + HF-LYP) [hartree]	$\langle S^2 \rangle$
singlet	-2756.15237298	0.3394
triplet	-2756.15211899	2.0013

DFT calculations.[9a, d, 11] The coupling constant $2J$ was evaluated by calculating the spin-projected energy difference between the singlet and triplet states. The obtained $2J$ value of $-45\ cm^{-1}$ compares well with the experimental value of $-77\ cm^{-1}$.

In summary, we have succeeded for the first time in determining the intermolecular exchange coupling for the light-induced RP in a crystal of *o*-Cl-HABI. This study shows that DFT methods can provide an accurate description even for the through-space coupling in RPs. A combination study of X-ray diffraction and DFT calculations for the light-induced RPs in HABI derivatives will lead to a better understanding of radical-pair chemistry and solid-state reactions.

Received: July 7, 2000 [Z15412]

[1] a) M. N. Paddon-Row, *Acc. Chem. Res.* **1994**, *27*, 18-25; b) A. Rajca, *Chem. Rev.* **1994**, *94*, 871-893; c) K. Olivier, *Molecular Magnetism*, VCH, New York, **1993**; d) J. S. Miller, A. J. Epstein, *Chem. Eng. News* **1995**, *73*(40), 30-41; e) C. B. Grissom, *Chem. Rev.* **1995**, *95*, 3-24.

[2] a) T. Hayashi, K. Maeda, *Bull. Chem. Soc. Jpn.* **1960**, *33*, 565-566; b) T. Hayashi, K. Maeda, *Bull. Chem. Soc. Jpn.* **1970**, *43*, 429-438; c) D. M. White, J. Sonnenberg, *J. Am. Chem. Soc.* **1966**, *88*, 3825-3829; d) H. Tanino, T. Kondo, K. Okada, T. Goto, *Bull. Chem. Soc. Jpn.* **1972**, *45*, 1474-1480, e) X.-Z. Qin, A. Liu, A. D. Trifunac, V. V. Krongauz, *J. Phys. Chem.* **1991**, *95*, 5822-5826; f) J. V. Caspar, I. V. Khudyakov, N. J. Turro, G. C. Weed, *Macromolecules*, **1995**, *28*, 636-641.

[3] M. Kawano, T. Sano, J. Abe, Y. Ohashi, *J. Am. Chem. Soc.* **1999**, *121*, 8106-8107.

[4] a) L. R. Ryzhkov, J. M. McBride, *J. Am. Chem. Soc.* **1997**, *119*, 4826-4833; b) J. M. McBride, B. E. Segmuller, M. D. Hollingsworth, D. E. Mills, B. A. Weber, *Science* **1986**, *234*, 830-835; c) D. A. Modarelli, P. M. Lahti, C. George, *J. Am. Chem. Soc.* **1991**, *113*, 6329-6330; d) H. L. Casal, D. Griller, R. J. Kolt, F. W. Hartstock, D. M. Northcott, J. M. Park, D. D. M. Wayner, *J. Phys. Chem.* **1989**, *93*, 1666-1670.

[5] K. Okada, K. Imamura, M. Oda, M. Kozaki, Y. Morimoto, K. Ishino, K. Tashiro, *Chem. Lett.* **1998**, 891-892.

[6] Y. Teki, I. Fujita, T. Takui, T. Kinoshita, K. Itoh, *J. Am. Chem. Soc.* **1994**, *116*, 11499-11505.

[7] S. S. Eaton, M. M. Kundalika, M. S. Bhimrao, G. R. Eaton, *J. Am. Chem. Soc.* **1983**, *105*, 6560.

[8] A fit of the plot for the thermal excited triplet state was performed with the following equation: $I = \text{const.} \times (1/T) \times \exp(\Delta E/RT)/[1 + 3 \times \exp(\Delta E/RT)]$, with I = EPR intensity, T = temperature, ΔE = energy difference, R = gas constant.

[9] a) E. Goldstein, B. Beno, K. N. Houk, *J. Am. Chem. Soc.* **1996**, *118*, 6063-6043; b) E. Ruiz, P. Alemany, S. Alvarez, J. Cano, *J. Am. Chem. Soc.* **1997**, *119*, 1297-1303; c) M. Abe, W. Adam, W. M. Nau, *J. Am. Chem. Soc.* **1998**, *120*, 11304-11310; d) J. H. Rodriguez, D. E. Wheeler, J. K. McCusker, *J. Am. Chem. Soc.* **1998**, *120*, 12051-12068.

[10] For the theoretical calculations, the molecular geometry of pair B was taken from the geometry determined in the X-ray diffraction study. The spin density of the triplet RP was calculated at the UB3LYP/6-31G* level. The zfs parameter D was derived from the following dipole-dipole equation: $|D/hc| = 3\beta g(\Sigma_i \Sigma_j \rho_i \rho_j / r_{ij}^3)$, where $3\beta = 1.299\ cm^{-1}Å^3$, g is the free electron g-value, β is the Bohr magneton value, ρ_i and ρ_j are interacting spin densities, and r_{ij} is the distance between them in angstroms.[9d] No interactions were allowed between densities that belong to the same lophyl radical.

[11] S. Yamanaka, T. Kawakami, H. Nagao, K. Yamaguchi, *Chem. Phys. Lett.* **1994**, *231*, 25-33.

The First Direct Determination of a Ligand Binding Constant in Protein Crystals

Su-ying Wu, Jacqueline Dornan, George Kontopidis, Paul Taylor, and Malcolm D. Walkinshaw*

In the field of protein-ligand interactions there have been many recent attempts to predict protein-ligand dissociation constants by correlating solution binding measurements with structural features from series of protein-ligand crystal complexes.[1-6] Indeed the whole structure-based design approach relies on the assumption that the thermodynamics of binding and the structures of bound ligands in the crystal will be similar to those in the biological environment. The results presented herein provide the first evidence that this is indeed the case.

Cyclophilins provide a good template for the study of ligand binding as there is relatively little change in their backbone conformation on binding.[7] To date the only structural data on ligand complexes is for cyclosporin derivatives or proline-containing peptides.[8-10] The isoform cyclophilin 3 from *Caenorhabditis elegans* (ceCyp3) was selected as the template

[*] Prof. Dr. M. D. Walkinshaw, Dr. S.-y. Wu, J. Dornan, Dr. G. Kontopidis, Dr. P. Taylor
Structural Biochemistry Group
Institute of Cell and Molecular Biology
The University of Edinburgh
Michael Swann Building, King's Buildings
Mayfield Road, Edinburgh, EH9 3JR (UK)
Fax: (+44) 131-650-7055
E-mail: M.Walkinshaw@ed.ac.uk

  1433-7851/01/4003-0582 $ 17.50+.50/0

for this binding study because of the availability of high symmetry, well-diffracting crystals which were amenable to ligand-soaking experiments. The enzymatic properties and the X-ray structure of the free form have been described.[11]

Solution binding and inhibition of ceCyp3 by the Ala-Pro ligand was measured by its ability to inhibit peptidylprolyl isomerase (PPIase) activity. The peptidylprolyl imide bond can adopt *cis* and *trans* conformations and the assay depends on the fact that chymotrypsin can only cleave the imide bond of the chromogenic prolyl substrate when it is in the *trans* conformation. The PPIase activity of cyclophilins catalyzes the rotation about this imide bond, thus speeding up the rate at which the *cis*–*trans* equilibrium is reached.[12, 13] Ala-Pro acts as a weak inhibitor of the PPIase activity, with a measured K_d value of 23.3 mM (Figure 1).

To measure the binding constant of the Ala-Pro ligand in the crystal a series of native ceCyp3 crystals were equilibrated against different concentrations of ligand (1.2–120 mM) and the individual X-ray structures refined. X-ray refinement statistics for the structures are given in Table 1. The refined structure of the crystal soaked in 120 mM Ala-Pro shows a well-defined ligand structure with seven hydrogen bonds involved in the ligand–protein binding (Figure 2b). The guanidino group of Arg62 adopts a very different orientation to that in the native structure and displaces water molecule W_d which is present in the native structure (Figure 2a). A further three water molecules (W_a, W_b, and W_c) present in the native structure are also displaced by the Ala-Pro ligand. All other residues around the binding site, with the exception of Arg62, adopt an identical conformation to that found in the native structure.

The (2Fo-Fc) electron density maps around the binding site of ceCyp3 complexed with different concentrations of Ala-Pro provide a graphic illustration of the conversion of the native to the ligand-bound structure (Figures 3a–g). The crystal soaked in a 120 mM solution of the ligand (Figure 3g) and the native crystal (Figure 3a) provide pictures of the two extremes. The five structures soaked with intermediate

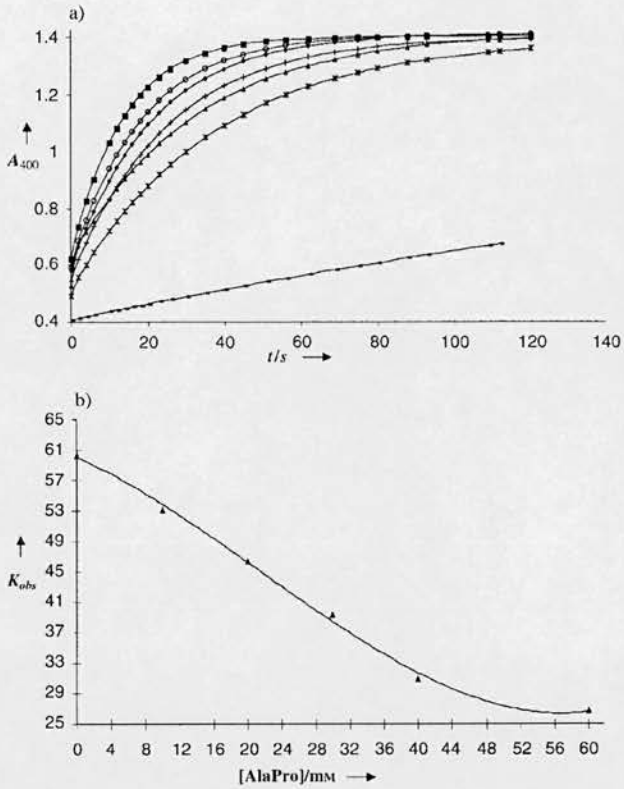


Figure 1. a) Inhibition of PPIase activity of ceCyp3 with increasing concentrations of Ala-Pro. Absorbance values were collected every 0.1 s over 2 min. The protein concentration was 20 nM and concentrations of Ala-Pro ranging between 0 and 60 mM were added. b) Determination of the K_d value for PPIase inhibition of ceCyp3 by Ala-Pro. Initial velocities from Figure 1a were used to determine k_{obsd} ($k_{obsd} = v/[S]$), where v is the velocity and [S] is the substrate concentration.[15] A plot of k_{obsd} against inhibitor concentration gives an Ala-Pro concentration of 23.3 mM for 50% inhibition. This IC_{50} value can be taken as an estimate for K_d by using $K_d = [E][L]/[EL]$ and at 50% inhibition $[E] = [EL] = 0.5[E]_T$ (where $[E_T]$ = total protein concentration, [E] = free protein concentration, [L] = free ligand concentration, and [EL] = protein concentration bound to ligand). With a large excess of ligand, the assumption $[L]_{50\%} = [L] \approx [L]_T$ can be made. At the point at which half maximal inhibition occurs $[EL] = [E]$ and $K_d = [L]_T$.

Table 1. Crystallographic statistics of the seven ceCyp3 structures.[a]

Data collection statistics	Native	1.2 mM Ala-Pro	6 mM Ala-Pro	18 mM Ala-Pro	36 mM Ala-Pro	60 mM Ala-Pro	120 mM Ala-Pro
space group, $P4_12_12$	$a=b=60.60$ Å	$a=b=60.87$ Å	$a=b=60.74$ Å	$a=b=61.03$ Å	$a=b=61.42$ Å	$a=b=61.05$ Å	$a=b=61.03$ Å
	$c=123.1$ Å	$c=122.99$ Å	$c=122.72$ Å	$c=122.7$ Å	$c=122.47$ Å	$c=122.67$ Å	$c=122.66$ Å
	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$	$\alpha=\beta=\gamma=90°$
resolution range	20–1.8 Å	25–2.1 Å	25–2.0 Å	20–1.9 Å	20–1.9 Å	25–2.0 Å	20–1.9 Å
number of reflections measured	145422	165474	189319	146663	224150	270345	152448
number of unique reflections	22053	14262	16311	19030	19185	16443	19230
completeness	97.0%	98.9%	99.7%	98.6%	99.5%	95.8%	97.6%
multiplicity	6.59	11.60	11.60	7.70	11.60	16.44	7.92
R_{merge} (outer shell)	8.3% (27.8%)	9.9% (36.2%)	8.1% (31.5%)	6.7% (34.6%)	7.2% (31.3%)	7.1% (33.2%)	7.9% (33.2%)
$I/\sigma(I)$ (outer shell)	11.7 (2.60)	12.50 (3.36)	14.88 (3.62)	13.88 (2.34)	14.89 (2.90)	14.44 (2.09)	12.85 (2.13)
R_{final} ($F>4\sigma$)	20.81% (19.69%)	17.65% (16.52%)	18.08% (17.09%)	20.93% (19.44%)	19.49% (18.32%)	22.56% (19.95%)	18.26% (17.3%)
R_{free} ($F>4\sigma$)	28.47% (26.61%)	25.94% (24.97%)	26.57% (24.45%)	29.58% (27.20%)	26.84% (24.62%)	29.16% (26.37%)	25.05% (23.77%)
rms deviations from ideal geometry							
bond lengths [Å]	0.031	0.030	0.030	0.030	0.031	0.030	0.030
angle distance [Å]	0.026	0.027	0.027	0.027	0.027	0.025	0.028
mean B-factor [Å^2]	30.37	33.70	31.93	32.73	30.95	35.84	31.36

[a] $R_{merge} = \Sigma_h |I - \langle I \rangle| / \Sigma_h |I|$, where $\langle I \rangle$ is the mean intensity of all observations of reflection $h = hkl$, $\sigma(I)$ is the standard deviation of the measured intensity, and R_{free} is calculated with 5% of the data omitted from refinement for cross-validation.

  1433-7851/01/4003-0583 $ 17.50+.50/0

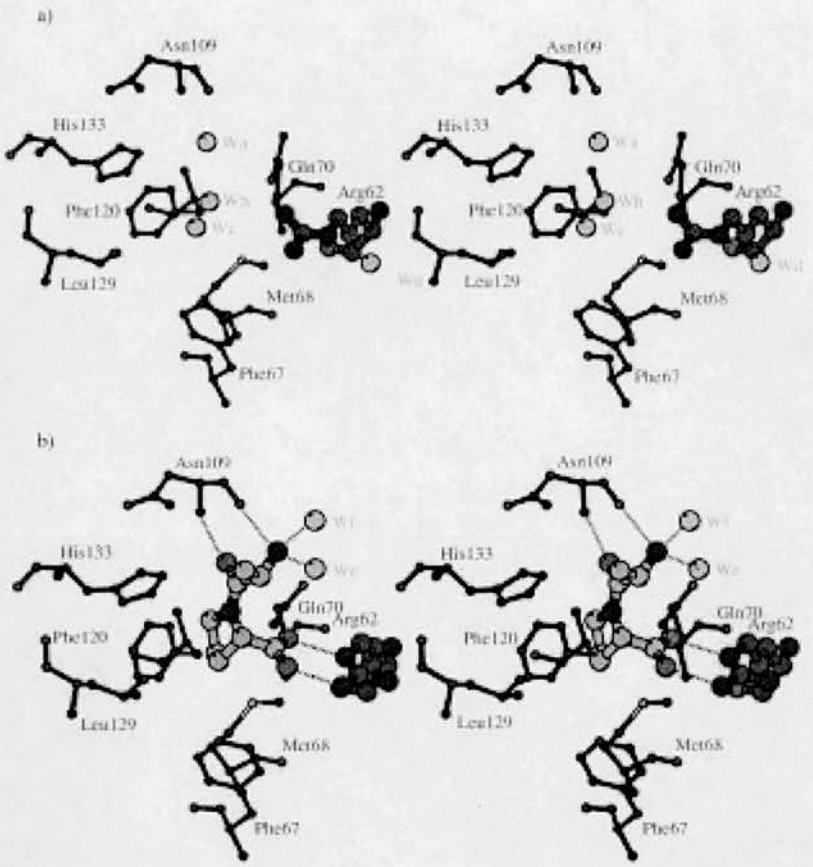


Figure 2. Active site of the structures with and without ligand. a) Stereoview of the active site of native (free) ceCyp3. The four water molecules in the active site ($W_a - W_d$) are coloured light gray. The side chain of Arg62 adopts a different conformation on ligand binding and is highlighted. b) Stereview of the active site of ceCyp3 after soaking in a 120 mM solution of Ala-Pro. The seven hydrogen bonds formed by the ligand to N/Asn109, NH_2/Arg62, W_e, O/Asn109, NE/Arg62, W_f, and NE2/Gln70 are shown as dotted black lines.

concentrations of Ala-Pro can be refined as mixtures of these two extreme structures. For example, the electron density maps of the crystals obtained in 18 mM and 36 mM solutions of the ligand (Figures 3d, e) show electron density that can clearly be interpreted as a mixture of the ligand-bound and free structures. The partial occupancy of the ligand (Q_l) is defined as the fraction of the protein molecules in the crystal occupied by the Ala-Pro ligand. Q_n is the fraction of protein molecules in the crystal in the native (free) state. As each molecule in the crystal must be in one of these states then $Q_n + Q_l = 1$. With this simple two-state model, occupancies can be constrained in the following way. The occupancy of the ligand (Q_l) is the same as the occupancy of the $Arg62_{(l)}$ in its ligand-bound conformation while the occupancy of the native

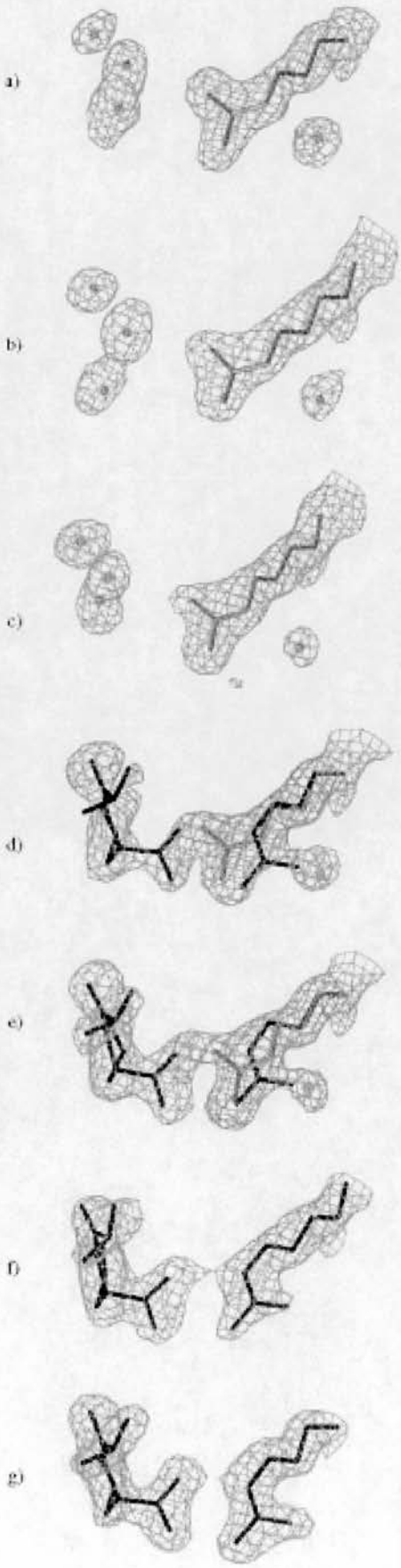


Figure 3. Difference electron density maps of crystals soaked with increasing ligand concentration. The (2Fo-Fc) maps highlighting the ligand, Arg62, and water molecules in the active site are contoured at 1.5 sigma using the program O.[16] a) The native ceCyp3 structure shows the density of the water molecules. b) and c) The crystals were soaked in 1.2 mM and 6 mM Ala-Pro, respectively. The electron density contoured at 1.5 sigma looks almost identical to the native structure. d) and e) The crystals were soaked in 18 mM and 36 mM Ala-Pro, respectively, and can be clearly interpreted as having features of both the native (light gray) and ligand-bound (black) forms. f) and g) The crystals were soaked in 60 and 120 mM Ala-Pro, respectively. The electron density contoured at 1.5 sigma shows only features corresponding with the fully ligand-bound form.

  1433-7851/01/4003-0584 $ 17.50+.50/0

structure (Q_n) is the occupancy of $Arg62_{(n)}$ in the native conformation, which must be the same as the occupancy of the water molecules W_a, W_b, W_c, and W_d. These occupancy constraints can be conveniently applied using the refinement programme SHELX97,[14] and the fractional occupancies refined. The occupancy is in fact only one parameter: $Q_l = 1 - Q_n$.

By regarding the crystal soaking experiment as a simple equilibrium in which each protein molecule in the crystal lattice can be bound or free, the fractional occupancy can be related to the dissociation constant in the following way:

$[PL] \rightarrow [P] + [L]$

$K_{cd} = [P][L]/[PL]$

$[P]_{total} = [PL] + [P]$

$[PL] = [P]_{total} Q_l$

$[P] = [P]_{total}(1 - Q_l)$

Thus,

$K_{cd} = \{[P]_{total}(1 - Q_l)\}[L]/[P]_{total} Q_l = (1 - Q_l)[L]/Q_l$

$Q_l/Q_n = [L]/K_{cd}$

where

K_{cd} is the crystal dissociation constant, [P] is the concentration of free protein, [L] is the concentration of the ligand (it is assumed that the concentration of ligand in the 10 μL drop is constant throughout the soaking experiment), [PL] is the concentration of protein with ligand, $[P]_{total}$ is the total concentration of protein, Q_l is the occupancy of ligand (fraction of the protein sites in the crystal occupied by ligand), Q_n is the fraction of the protein sites in the free state ($=(1 - Q_l)$).

A value of 26.8 mM was obtained for K_{cd} as calculated from the slope of the plot of (Q_l/Q_n) versus [L] (Figure 4b). This value is in close agreement with the inhibition constant of 23.3 mM measured for the Ala-Pro ligand in solution using the PPIase assay. It is noteworthy that the quality and information content of the X-ray data is sufficient to accurately refine the partial occupancy of the ligand-bound structures even when visual inspection of the electron density maps suggests the electron density is dominated by one of the (native or ligand-bound) forms. In general, interpreting a structure with such low partial occupancy would be almost impossible using conventional map-fitting methods without a well-refined native structure being available. Indeed inspection of the series of electron density maps shown in Figure 3 suggests that, without prior knowledge, side-chain conformations with occupancies less than about 20% would not be readily apparent in electron density maps.

In protein–ligand complexes it can frequently be difficult to distinguish water molecules or disordered water molecules from a weakly bound ligand. A helpful feature of this system is the movement of the Arg62 side chain on ligand binding, as this gives an excellent double-check on the ligand occupancy. In a separate experiment we also showed that diffusion and binding of the ligand in this system took place on a timescale of seconds. A crystal soaked in 120 mM ligand solution for 60 seconds was flash frozen and the refined structure shown to be identical to that from a crystal soaked for 17 hours. The kinetics of ligand–protein binding is likely to vary somewhat from crystal to crystal depending on the crystal packing, solvent channel size, and ligand size. However, we suggest that

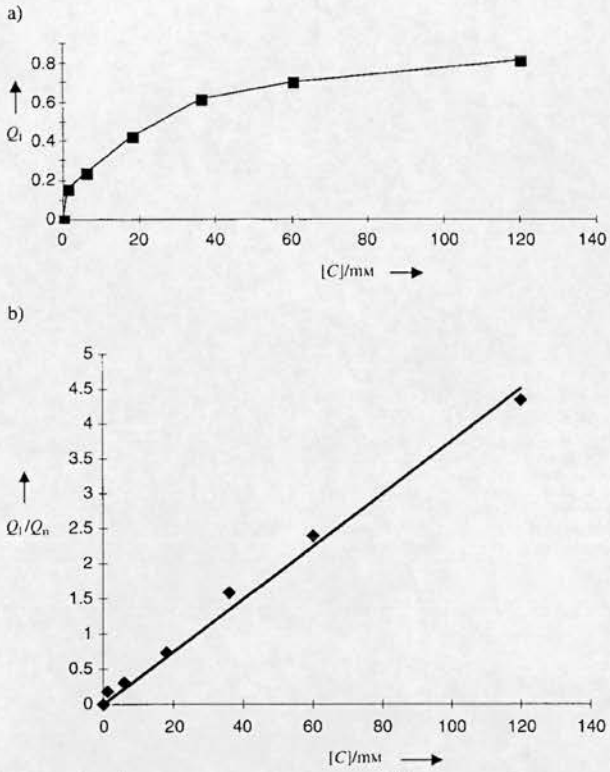


Figure 4. a) Refined occupancy for the seven different crystal structures. Each point on the curve represents the occupancy of the Ala-Pro ligand as determined from an X-ray refinement of a crystal soaked in a given concentration of Ala-Pro.[17] The ligand concentrations used in the soaking experiments were 0, 1.2, 6, 18, 36, 60, and 120 mM. The corresponding occupancies obtained from crystallographic refinement were 0, 0.155, 0.238, 0.424, 0.613, 0.705, and 0.812. b) Plot of ligand concentration against Q_l/Q_n. Q_l: occupancy of ligand (fraction of the protein sites in the crystal occupied by ligand), Q_n: fraction of the protein sites in the free state ($=(1 - Q_l)$). The ligand occupancies of the different crystals were measured by crystallographic refinement as described in the text. The experimental data fit the theoretical line with a high R^2 value of 0.990 and a slope $1/K_{cd} = 0.0373$. This gives a value for K_{cd} (the crystal dissociation constant) of 26.8 mM.

for a large number of proteins, this crystallographic method of measuring binding constants could be developed into a relatively efficient and routine technique.

Experimental Section

Crystallization and data collection: CeCyp3 was crystallized at 16 °C by the hanging drop vapour-diffusion method using 29–31% poly(ethylene glycol) methyl ether 5000 (MPEG 5000) as precipitant.[11] Native ceCyp3 crystals (approximately 0.2 × 0.3 × 0.3 mm) were soaked in six 10 μL drops of well solution containing the Ala-Pro dipeptide at 1.2, 6, 18, 36, 60, and 120 mM, respectively. Crystals were soaked for between 17 and 27 hours.

The composition of the cryoprotectant was glycerol (0.2 mL), MPEG 5000 (0.7 mL), 1.0 M sodium citrate/citric acid (0.1 mL, pH 5.6). Crystals were mounted in a 0.2–0.3-mm cryoloop (Hampton Research, Inc.), dipped in cryoprotectant for a few seconds, and frozen by plunging into liquid nitrogen. The frozen crystal was then transferred to a magnetic goniometer head in a stream of nitrogen at 100 K. All diffraction data were collected on a 300-mm MarResearch imaging plate system mounted on an ENRAF Nonius FR571 rotating anode generator operating at 40 kV, 80 mA, and producing $Cu_{K\alpha}$ radiation from a graphite crystal monochromator.

Refinement: The ceCyp3 native structure was used as a starting model. Positional and B-factor refinement were performed using SHELX97

  1433-7851/01/4003-0585 $ 17.50+.50/0

(Table 1). The coordinates of $Arg62_{(n)}$, $Arg62_{(l)}$, Ala-Pro, W_a, W_b, W_c, and W_d were fixed during occupancy refinement. $Arg62_{(n)}$ represents the Arg62 conformation without Ala-Pro bound and $Arg62_{(l)}$ represents the Arg62 conformation with Ala-Pro bound to ceCyp3. The B-factors of $Arg62_{(n)}$, W_a, W_b, W_c, and W_d were fixed to the same B-factor values as those of the native structure. Individual atomic B-factors for $Arg62_{(l)}$ and Ala-Pro were refined together with occupancy (Figure 4). The restraints applied in the occupancy refinement are summarized below:

$Q_{Arg62(n)} + Q_{Arg62(l)} = 1$

$Q_{Arg62(l)} = Q_l$

$Q_{Arg62(n)} = Q_{Wa} = Q_{Wb} = Q_{Wc} = Q_{Wd}$

where

$Q_{Arg62(n)}$: is the occupancy of the Arg62 conformation with no Ala-Pro binding, $Q_{Arg62(l)}$: is the occupancy of the Arg62 conformation with Ala-Pro binding, and Q_l: is the occupancy of Ala-Pro.

PPIase assay:[13] α-chymotrypsin selectively hydrolyzes the C-terminal *p*-nitroanilide bond of the substrate in the *trans* X-Pro conformer only. This hydrolysis releases the chromophore 4-nitroaniline, the accumulation of which is recorded by measuring the absorbance at 400 nm as a function of time. Substrate (*N*-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide, Bachem AG) was dissolved in LiCl/2,2,2-trifluoroethanol (LiCL/TFE) to give a stock solution of 100 mM. The experiment took place at 4 °C. Constant temperature was maintained within the cuvette by a Peltier (PTP-1) temperature control unit. A Perkin–Elmer UV/Vis Lambda 20 spectophotometer was used.

Proteins: ceCyp3 solution was freshly prepared before the experiment from frozen stock solution, at the appropriate concentration, by dilution in buffer 50 mM 2-[4-[2hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES), 100 mM NaCl, pH 8.0 (buffer A).

α-chymotrypsin (Sigma): In a typical experiment 10 µL of 20 nM ceCyp3 was made up to 870 µL with buffer A and the appropriate volume of Ala-Pro in a 1-mL cuvette. The cuvette was then preincubated for 30 min on ice. Immediately before the assay, 100 µL of chymotrypsin solution (50 mg mL^{-1} in 10 mM HCl) was added, followed by 30 µL of a 3.7 µM stock solution of Suc-Ala-Ala-Pro-PNA in LiCl (470 mM)/TFE. The reaction progress was monitored by the absorbance change at 400 nm that accompanies the hydrolysis of the amide bond and the release of 4-nitroaniline.

Received: September 8, 2000 [Z15779]

[1] B. M. Baker, K. P. Murphy, *Methods Enzymol.* **1998**, *295*, 294–315.
[2] B. K. Shoichet, A. R. Leach, I. D. Kuntz, *Proteins* **1999**, *34*, 4–16.
[3] G. M. Verkhivker, P. A. Rejto, D. Bouzida, S. Arthurs, A. B. Colson, S. T. Freer, D. K. Gehlhaar, V. Larson, B. A. Luty, T. Marrone, P. W. Rose, *J. Mol. Recognit.* **1999**, *12*, 371–389.
[4] M. Schapira, M. Totrov, R. Abagyan, *J. Mol. Recognit.* **1999**, *12*, 177–190.
[5] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, R. P. Mee, *J. Comput. Aided Mol. Des.* **1997**, 11, 425–445.
[6] H. J. Bohm, *J. Comput. Aided Mol. Des.* **1998**, *12*, 309–323.
[7] P. Taylor, H. Husi, G. Kontopidis, M. D. Walkinshaw, *Prog. Biophys. Mol. Biol.* **1997**, *67*, 155–181.
[8] J. Kallen, M. D. Walkinshaw, *FEBS Lett.* **1992**, *300*, 286–290.
[9] J. Kallen, V. Mikol, P. Taylor, M. D. Walkinshaw, *J. Mol. Biol.* **1998**, *283*, 435–449.
[10] Y. D. Zhao, H. M. Ke, *Biochemistry* **1996**, *35*, 7362–7368.
[11] J. Dornan, A. P. Page, P. Taylor, S. Y. Wu, A. D. Winter, H. Husi, M. D. Walkinshaw, *J. Biol. Chem.* **1999**, *274*, 34877–34883.
[12] D. Kern, G. Kern, G. Scherer, G. Fischer, T. Drakenberg, *Biochemistry* **1995**, *34*, 13594–13602.
[13] J. L. Kofron, P. Kuzmic, V. Kishore, E. Colonbonilla, D. H. Rich, *Biochemistry* **1991**, *30*, 6127–6134.
[14] G. M. Sheldrick, SHELX97, University of Göttingen, **1997**.
[15] J. Liu, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4028–4032.
[16] T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. Sect. A* **1991**, *47*, 110–119.
[17] Crystallographic data for the structures reported in this paper have been deposited with the Protein Data Bank as supplementary publication no. PDB-1E8K (see http://www.rcsb.org/pdb/index.html).

A Heterogeneous *cis*-Dihydroxylation Catalyst with Stable, Site-Isolated Osmium–Diolate Reaction Centers**

An Severeyns, Dirk E. De Vos, Lucien Fiermans, Francis Verpoort, Piet J. Grobet, and Pierre A. Jacobs*

Osmium tetroxide is by far the most versatile catalyst for *cis*-dihydroxylation (DH) of double bonds.[1, 2] When homogeneous catalysts are used, free OsO_4 is always present in some step of the catalytic cycle, and the high toxicity and volatility of OsO_4 have hitherto obstructed industrial application. Previous attempts to immobilize OsO_4 used polymers, for example, with coordination of OsO_4 on polyvinylpyridine.[3, 4] However, hydrolysis of the intermediate Os^{VI} diolate complex requires that Os is detached from the polymeric Lewis base,[5] and this implies an inherent liability to Os leaching. Similarly, reports on immobilized alkaloids for asymmetric DH mention that Os leaching necessitates Os supplementation in subsequent runs.[6] In another attempt, OsO_4 was entrapped in polystyrene microspheres, but the mechanism by which OsO_4 is retained within the polymer is not understood.[7] Herein we report a solid with Os^{VIII} type reactivity, and with a persistent bond between Os and the support. Rigorous heterogeneity tests and reactions with 12 olefins substantiate the value of the new Os catalyst.

Our approach is rooted in the mechanism of the *cis*-dihydroxylation, which comprises two stages: 1) attack of the Os^{VIII} *cis*-dioxo complex on the olefin (osmylation), 2) reoxidation of Os^{VI} to Os^{VIII} and hydrolytic release of the diol. Two points are particularly relevant. First, if the hydrolytic conditions are not too drastic, *tetrasubstituted* olefins are not converted into *cis*-diols.[8, 9] These olefins are smoothly osmylated to an osmate(VI) ester, but the rate of subsequent hydrolysis is zero (0% yield for a tetrasubstituted olefin vs. 83% for a trisubstituted olefin, ref. [8]). Second, an Os^{VI} monodiolate complex can be reoxidized to *cis*-dioxo Os^{VIII} without release of the diol; subsequent addition of a second olefin results in an Os bisdiolate complex.[10] These two properties make it possible to immobilize a catalytically active Os compound by the addition of OsO_4 to a tetrasubstituted olefin that is covalently linked to a silica support (**1a**, Scheme 1). The tetrasubstituted diolate ester (**1b**) which is

[*] Prof. Dr. Ir. P. A. Jacobs, Ir. A. Severeyns, Prof. Dr. Ir. D. E. De Vos, Prof. Dr. P. J. Grobet
Center for Surface Chemistry and Catalysis
Katholieke Universiteit Leuven
Kardinaal Mercierlaan 92, 3001 Heverlee (Belgium)
Fax: (+32) 16-32-1998
E-mail: pierre.jacobs@agr.kuleuven.ac.be
Prof. Dr. L. Fiermans
Laboratory for Crystallography and Study of Solid Matter
Universiteit Gent (Belgium)
Prof. Dr. F. Verpoort
Deptment of Inorganic and Physical Chemistry
Universiteit Gent (Belgium)

[**] This work was supported by the Belgian Federal Government in the frame of an Interuniversitary Attraction Pole on Supramolecular Catalysis. We are indebted to FWO for fellowships (A.S., D.D.V.) and for a research grant (D.D.V., F.V., P.A.J.).

  1433-7851/01/4003-0586 $ 17.50+.50/0

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, characterization and crystallization in two crystal forms of bovine cyclophilin 40

Jacqueline Dornan,[a] Paul Taylor,[a] Amerigo Carrello,[b,c] Rodney F. Minchin,[c] Thomas Ratajczak[b,c] and Malcolm D. Walkinshaw[a]*

[a]Structural Biochemistry Group, Institute of Cell and Molecular Biology, The University of Edinburgh, Michael Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland, [b]Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Hospital Avenue, Nedlands WA 6009, Australia, and [c]Department of Pharmacology, University of Western Australia, The Queen Elizabeth II Medical Centre, Nedlands WA 6009, Australia

Correspondence e-mail: m.walkinshaw@ed.ac.uk

Received 7 December 1998
Accepted 22 February 1999

The purification and crystallization of two different crystal forms of the two-domain protein bovine cyclophilin 40 is reported. Tetragonal crystals grown in methyl pentanediol belong to space group $P4_222$ with unit-cell parameters a = 94.5, c = 118.3 Å. Long thin needles grown from PEG belong to space group $C2$ with unit-cell parameters a = 125.71, b = 47.3, c = 74.6 Å, β = 93.90°. The N-terminal 170 amino acids have significant homology with the well characterized human cyclophilin A. The C-terminal domain is largely made up of three copies of the tetratricopeptide repeat motif thought to be involved in mediating protein–protein interactions. Cyclophilins are frequently found as domains in larger multidomain proteins. To date, only X-ray structures of single-domain cyclophilins have been reported, and this work provides the first example of the purification and crystallization of a larger protein containing a cyclophilin domain.

1. Introduction

Cyclophilins are phylogenetically highly conserved proteins which act as peptidyl–prolyl *cis–trans* isomerases (PPIases) and are involved in protein folding/refolding as well as protein–protein interactions (Galat & Metcalfe, 1995). The first member of this expanding family to be discovered was the 18 kDa human cyclophilin A (hCyP18) which was identified by its binding to the fungal peptide immunosuppressant cyclosporin A (CsA) (Handschumacher *et al.*, 1984). The X-ray structures of cyclophilin A and of its complex with CsA have been determined (Kallen *et al.*, 1991; Pflugl *et al.*, 1993) and show the eight-stranded barrel structure conserved in all cyclophilin domains.

Larger multidomain proteins such as CyP150 from natural killer cells (Anderson *et al.*, 1993) and RanBP2 (Yokoyama *et al.*, 1995) also contain single cyclophilin domains. The two-domain cyclophilin 40 (CyP40) protein family has been found in mammals (Kieffer *et al.*, 1992 1993; Ratajczak *et al.*, 1993) and yeast (Weisman *et al.*, 1996; Duina *et al.*, 1996), and was originally isolated and characterized in association with the unactivated non-DNA binding form of the oestrogen receptor (Ratajczak *et al.*, 1993). CyP40 is also present as one of several components of mature unactivated complexes of the glucocorticoid and progesterone receptors (reviewed in Pratt & Toft, 1997). The major chaperones, hsp90 and hsp70, are included as non-hormone-binding components within these steroid-receptor complexes, together with p23 and the immunophilins FKBP51 and FKBP52 (Pratt & Toft, 1997), which are cellular targets of the immunosuppressant FK506 (Galat & Metcalfe, 1995). CyP40, FKBP51 and FKBP52 belong to a distinct class of proteins which display an N-terminal immunophilin-like domain with overlapping regions for immunosuppressant drug interaction and isomerase activity, together with a conserved C-terminal tetratricopeptide repeat (TPR) domain proposed to mediate protein–protein interaction (Ratajczak *et al.*, 1993; Nair *et al.*, 1997). Through this conserved C-terminal region, these immunophilins target a common interaction site in hsp90, resulting in the formation of separate immunophilin–hsp90–steroid-receptor complexes (Nair *et al.*, 1997; Radanyi *et al.*, 1994; Hoffmann & Handschumacher, 1995; Ratajczak & Carrello, 1996; Young *et al.*, 1998; Carrello *et al.*, 1999). CyP40 and its partner immunophilins appear to have a coordinate role with hsp90 in the functional control of receptor activity (Pratt & Toft, 1997).

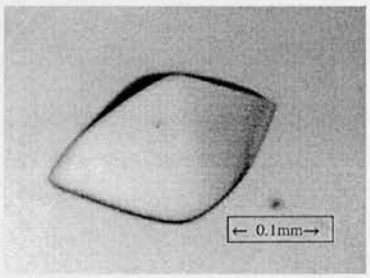
In vitro folding assays have shown that Cyp40 and FKBP52 can function as molecular chaperones, similar to hsp90 and hsp70, by holding substrate proteins in a partially folded conformation (Freeman *et al.*, 1996; Bose *et al.*, 1996). The expression of these immunophilins is ubiquitous (Ratajczak *et al.*, 1993; Kieffer *et al.*, 1993; Nair *et al.*, 1997) and there is evidence for their association with hsp90 independently of steroid receptors (Pratt & Toft, 1997). Cpr6 and Cpr7 are the only *Saccharomyces cerevisiae* homologues of Cyp40, and both associate with hsp90 (Duina *et al.*, 1996). Mutations which decrease the level of hsp90 expression, coupled with a deletion of Cpr7, cause a

significant reduction in the rate of cell growth (Duina *et al.*, 1996). The involvement of wis2, the corresponding CyP40 homolog in *Schizosaccharomyces pombe*, in regulating the G2-mitosis transition (Weisman *et al.*, 1996) is consistent with this observation. Hsp90 in association with Cpr7 has also been shown to participate in the negative regulation of the heat-shock response in *S. cerevisiae* (Duina *et al.*, 1998). Taken together, these results suggest that Cyp40 plays a major role in hsp90-mediated signal-transduction pathways.

2. Materials and methods

2.1. Expression, purification and characterization

Expression plasmids for untagged wild-type bovine CyP40 (bCyP40) and the cyclophilin fused to glutathione S-transferase (GST) have been described (Ratajczak & Carrello, 1996; Carrello *et al.*, 1999).



(*a*)

(*b*)

Figure 1
Pictures of the monoclinic and tetragonal crystal forms of Cyp40.

GST-fused recombinant cyclophilin 40 (GST-bCyP40 WT; Ratajczak & Carrello, 1996) was expressed following a 4 h induction with 0.4 m*M* IPTG (ICN/FLOW). Cell pellets were frozen at 253 K, thawed slowly on ice and solubilized in lysis buffer [20 m*M* Tris buffer pH 7.8 containing 150 m*M* NaCl, 2 m*M* EDTA, 5 m*M* DTT, 1%(*v*/*v*) Triton X-100 and 5 m*M* benzamidine]. Lysozyme (ICN/FLOW) was added to 0.1%(*w*/*v*) and the cell suspension incubated on ice for 1 h. An equal volume of distilled water at 277 K was added and the cell suspension was mixed and centrifuged at 48400*g* for 60 min. The resulting supernatant was frozen at 253 K, thawed on ice and filtered (0.45 µm) prior to purification. The protein extract was purified to near homogeneity using four purification steps. (i) Glutathione Sepharose (Pharmacia) affinity chromatography: protein was applied in buffer *A* (50 m*M* Tris–HCl pH 7.5 buffer) and GST-tagged protein was eluted with buffer *A* containing reduced 25 m*M* glutathione (Sigma). (ii) Anion-exchange chromatography: protein from the glutathione Sepharose chromatography was applied to Resource-Q resin (Pharmacia) in buffer *A*. Elution with a linear gradient from buffer *A* to buffer *A* plus 1 *M* NaCl gave peak fractions which were pooled, concentrated, quantitated for protein concentration by the Bradford method and dialysed overnight against buffer *A*. (iii) Thrombin cleavage: GST fusion protein was cleaved with purified thrombin at a ratio of 1:100 for 90 min at room temperature. (iv) Cleaved protein was subjected to a second round of chromatography on Resource-Q resin as already described, followed by a final affinity chromatography step on glutathione Sepharose.

Induction and harvesting of the non-fusion version of the protein was essentially as described above for the GST-bCyp40 WT protein. Purification to near homogeneity was achieved in two ion-exchange steps. The cell extract was applied to Q Sepharose (Amersham Pharmacia Biotech) in buffer *A*, the column was washed to baseline and then developed with a gradient of buffer *A* plus 1 *M* NaCl. Protein was eluted with 150 m*M* NaCl and fractions were analysed on 12% SDS–PAGE gels followed by staining with Coomassie Blue. Fractions containing bCyP40 were pooled and dialysed overnight at 277 K against 20 m*M* Tris pH 8.0 buffer (buffer *B*). The protein sample was applied to Mono-Q resin (PAB) in buffer *B*, and retained protein was eluted with a gradient of buffer *B* plus 0.5 *M* NaCl. Positive fractions were pooled and concentrated. Dynamic light-scattering analysis of the purified protein (DynaPro-801 with Microsampler attatchment, Protein Solutions, Inc.) showed it to be essentially monodisperse and stable, with an estimated molecular weight of 43 kDa.

The protein was further characterized by electrospray ionization mass spectrometry on a Micromass Platform II spectrometer. The sample was introduced into the instrument stream *via* an infusion pump and the *M*/*Z* spectrum was deconvoluted using the *MaxEnt* algorithm.

2.2. Crystallization

The hanging-drop vapour-diffusion method was used to grow crystals in 24-well Linbro plates. In all trials, crystals could only be obtained at 277 K.

Tetragonal crystals were grown by equilibrating a 4 µl drop of protein solution against 1 ml of well solution. The well solution consisted of 470–490 µl MPD, 100 µl 1 *M* imidazole pH 6.5, 200–360 µl 50%(*v*/*v*) glycerol solution and 50–230 µl water. The drop consisted of 2 µl protein solution and 2 µl well solution. The protein solution consisted of protein at between 10 and 80 mg ml^{-1} in 20 m*M* Tris buffer pH 8 containing 100 m*M* NaCl.

The morphology of the crystals in most drops was poorly defined and teardrop-shaped crystals grew with dimensions of ~0.05 mm. Addition of 1% glycerol and 5–10 m*M* DTT were found to improve crystal size and morphology (Fig. 1*a*), with crystals growing to 0.2 × 0.15 × 0.15 mm. The unit-cell dimensions are a = 94.5, c = 118.3 Å. A complete data set to a resolution of 2.6 Å has been collected on an Enraf–Nonius rotating-anode X-ray generator with a MAR Research 300 image plate.

Monoclinic needles were obtained using 12%(*v*/*v*) PEG as precipitant. The 1 ml well solution was composed of 240 µl of a 50% solution of PEG 2000, 100 µl 1 *M* cacodylate, 50 µl 1 *M* HCl, 200 µl 50% glycerol solution and 410 µl water. The 4 µl hanging drop consisted of 2 µl well solution plus 2 µl protein at 10–80 mg ml^{-1} in 20 m*M* Tris pH 8 buffer containing 100 m*M* NaCl. Very long thin plate-like needles grew within hours with typical dimensions 2 × 0.01 × 0.005 mm. These crystals are monoclinic with space group *C*2. Unit-cell parameters are a = 125.71, b = 47.3, c = 74.6 Å, β = 93.90°. These thin needle crystals typically diffract to a resolution of 3.6 Å on a rotating-anode X-ray source, and a complete data set has been collected.

```
                1                                                   50
CYPH_HUMAN      .......... ...VNPTVFF DIAVDGEPLG RVSFELFADK VPKTAENFRA
CYP3_CAEEL      .......... ..MSRSKVFF DITIGGKASG RIVMELYDDV VPKTAGNFRA
CYP4_HUMAN      MSHPSPQAKP SNPSNPRVFF DVDIGGERVG RIVLELFADI VPKTAENFRA
CYP4_BOVIN      MSHPSPQAKP SNPSNPRVFF DVDIGGERVG RIVLELFADI VPKTAENFRA
 Consensus      .......... .......vff d....g...g r...el..d. vpkta.nfra

                51                                                  100
CYPH_HUMAN      LSTGEKGFG. .......YKG SCFHRIIPGF MCQGGDFTRH NGTGGKSIYG
CYP3_CAEEL      LCTGENGIGK S.GKPLHFKG SKFHRIIPNF MIQGGDFTRG NGTGGESIYG
CYP4_HUMAN      LCTGEKGIGH TTGKPLHFKG CPFHRIIKKF MIQGGDFSNQ NGTGGESIYG
CYP4_BOVIN      LCTGEKGIGP TTGKPLHFKG CPFHRIIKKF MIQGGDFSNQ NGTGGESIYG
 Consensus      l.tge.g.g. ........kg ..fhrii..f m.qggdf... ngtgg.siyg

                101                                                 150
CYPH_HUMAN      EKFEDENFIL KHTGPGILSM ANAGPNTNGS QFFICTAKTE WLDGKHVVFG
CYP3_CAEEL      EKFPDENFKE KHTGPGVLSM ANAGPNTNGS QFFLCTVKTE WLDGKHVVFG
CYP4_HUMAN      EKFEDENFHY KHDREGLLSM ANAGRNTNGS QFFITTVPTP HLDGKHVVFG
CYP4_BOVIN      EKFEDENFHY KHDKEGLLSM ANAGSNTNGS QFFITTVPTP HLDGKHVVFG
 Consensus      ekf.denf.. kh...g.lsm anag.ntngs qff..t..t. .ldgkhvvfg

                151                                                 200
CYPH_HUMAN      KVKEGMNIVE AMERFGSRNG KTSKKITIAD CGQLE..... ..........
CYP3_CAEEL      RVVEGLDVVK AVESNGSQSG KPVKDCMIAD CGQLKA.... ..........
CYP4_HUMAN      QVIKGIGVAR ILENVEVKGE KPAKLCVIAE CGELKEGDDG GIFPKDGSGD
CYP4_BOVIN      QVIKGMGVAK ILENVEVKGE KPAKLCVIAE CGELKEGDDW GIFPKDGSGD
PPP5_HUMAN      .......... .......... .......... .......... .....MAMAE
 Consensus      .v..g..... ..e..... k..k...ia. cg.l...... ..........

                201                                                 250
CYPH_HUMAN      .......... .......... .......... .......... ..........
CYP3_CAEEL      .......... .......... .......... .......... ..........
CYP4_HUMAN      SHPDFPEDAD IDLKDVDKIL LITEDLKNIG NTFFKSQNWE MAIKKYAEVL
CYP4_BOVIN      SHPDFPEDAD VDLKDVDKIL LISEDLKNIG NTFFKSQNWE MAIKKYTKVL
PPP5_HUMAN      GERTECAEPP RDEPPADGAL KRAEELKTQA NDYFKAKDYE NAIKFYSQAI
 Consensus      .......... .d....d..l ...e.lk... n..fk....e .aik.y....

                251                                                 300
CYPH_HUMAN      .......... .......... .......... .......... ..........
CYP3_CAEEL      .......... .......... .......... .......... ..........
CYP4_HUMAN      RYVDSSKAVI ETADRAKLQP IALSCVLNIG ACKLKMSNWQ GAIDSCLEAL
CYP4_BOVIN      RYVEGSRAAA EDADGAKLQP VALSCVLNIG ACKLKMSDWQ GAVDSCLEAL
PPP5_HUMAN      .......... ......ELNP SNAIYYGNRS LAYLRTECYG YALGDATRAI
 Consensus      .......... .......l.p .......n.. ...l...... .a......a.

                301                                                 350
CYPH_HUMAN      .......... .......... .......... .......... ..........
CYP3_CAEEL      .......... .......... .......... .......... ..........
CYP4_HUMAN      ELDPSNTKAL YRRAQGWQGL KEYDQALADL KKAQGIAPED KAIQAELLKV
CYP4_BOVIN      EIDPSNTKAL YRRAQGWQGL KEYDQALADL KKAQEIAPED KAIQAELLKV
PPP5_HUMAN      ELDKKYIKGY YRRAASNMAL GKFRAALRDY ETVVKVKPHD KDAKMKYQEC
 Consensus      e.d....k.. yrra.....l .....al.d. .......p.d k.........

                351               370
CYPH_HUMAN      .......... ..........
CYP3_CAEEL      ..........: ..........
CYP4_HUMAN      KQKIKAQKDK EKAVYAKMFA
CYP4_BOVIN      KQKIKAQKDK EKAAYAKMFA
PPP5_HUMAN      NKIVK.QKAF ERAIAGDEHK....
 Consensus      ....k.qk.. e.........
```

(*a*)

```
CYP4_BOVIN          LISEDLKNIGNTFFKSQNWEMAIKKYTKVLRYVEGSRAAAEDADGAKLQP
PPP5_HUMAN          KRAEELKTQANDYFKAKDYENAIKFYSQAIELNP..... ......

CYP4_BOVIN          VALSCVLNIGACKLKMSDWQGAVDSCLEALEIDP
PPP5_HUMAN_         SNAIYYGNRSLAYLRTECYGYALGDATRAIELDK

CYP4_BOVIN          SNTKALYRRAQGWQGLKEYDQALADLKKAQEIAPEDKAIQAELLKVKQKIKAQKDKEKAAYAKMFA
PPP5_HUMAN          KYIKGYYRRAASNMALGKFRAALRDYETVVKVKPHDKDAKMKYQECNKIVK QKAFERAIAGDEHK.....
```

(*b*)

Figure 2
(*a*) Sequence alignment of human cyclophilin A (CYPH_HUMAN), nematode cyclophilin 3 (CYP3_CAEEL), human and bovine cyclophilin 40 (CYP4_HUMAN, CYP4_BOVIN) and human protein phosphatase 5 (PPP5_HUMAN). Sequences from the Swiss–Prot database were aligned using the program *MultAlign* and represent the entire sequence of the cyclophilins aligned against the N-terminal 150 amino acids from PP5. The TPRs of PP5 are shown in bold. Numbering applies to the amino-acid sequence of bovine CyP40. (*b*) Alignment of the three TPR motifs of bCyP40 against the N-terminal 150 TPR domain of PP5. The six helices associated with the three TPR motifs determined in the the X-ray structure of PP5 (Das *et al.*, 1998) are underlined. The C-terminal extension of 28 amino acids is the predicted helical calmodulin-binding site. The program *MultAlign* (Corpet, 1988) was used for amino-acid alignment.

3. Discussion

The amino-acid sequence for bCyP40 is shown in Fig. 2(*a*) and shares 94% identity with the human homologue. Mass spectrometry gives a molecular mass of 40500 Da. The calculated molecular mass of the sequence shown in Fig. 2 is 40620 Da, which suggests that the N-terminal methionine residue is removed post-translationally. The calculated molecular mass of the sequence omitting the N-terminal methionine is 40489 Da.

The cyclophilin domain, consisting of the N-terminal 187 residues, is very well conserved between the bovine and human proteins, with differences in only five residues. The N-terminal domain is 61% identical to hCyP18 and shows an even better homology (64%) with the so-called divergent class of cyclophilins (Taylor *et al.*, 1998), which have an additional seven amino-acid loop inserted at positions 60–67 (Fig. 2*a*). Another feature of this divergent class of cyclophilins is that the active-site tryptophan, shown to be important in binding CsA in hCyP18 (Pflugl *et al.*, 1994), is replaced by a histidine. This change may explain the reduction of the IC_{50} values for CsA from 20 n*M* for hCyP18 to 300 n*M* for hCyP40 (Kieffer *et al.*, 1992). The CyP40/CsA/calcineurin K_i was 320 n*M*, compared with the CyP18/CsA/calcineurin K_i of 195 n*M*, and CyP40 was found to compete with CyP18 for binding to calcineurin in the presence of CsA (Kieffer *et al.*, 1993).

The C-terminal 150 amino acids (residues 188–359) of hCyP40, comprising the TPR and putative calmodulin-binding domains, is 31% homologous with the corresponding region of FKBP59. An X-ray structure is available for the TPR domain of the human protein PP5 (Das *et al.*, 1998) which shows that each TPR motif consists of two helices of about 14 amino acids which are linked by a tight turn. The predicted secondary structure of the C-terminal CyP40 domain is shown in Fig. 2(*b*), along with a sequence alignment between human phosphatase and bCyP40. Helical regions of the predicted TPR motifs are also shown.

The PP5 crystal structure (Das *et al.*, 1998) shows that the putative calmodulin-binding domain of CyP40 consists of an α-helix approximately 30 amino acids in length. The secondary-structure prediction (Rost *et al.*, 1994) of the CyP40 sequence gives a strongly predicted C-terminal helix of 28 amino acids in length. This corresponds to a length of 42 Å and is close to the conserved *b* cell dimension of the monoclinic needle form of CyP40. The *b* axis also corresponds to the

long and fast-growing needle axis, and it is reasonable to suggest that it is this calmodulin-binding helix which determines the length of the *b* axis and also the kinetics and morphology of the monoclinic needle crystals.

We thank Violet R. Anderson for the mass spectrometry.

References

Anderson, S. K., Gallinger, S., Roder, J., Frey, J., Young, H. A. & Ortaldo, J. R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 542–546.

Bose, S., Weikl, T., Bugl, H. & Buchner, J. (1996). *Science*, **274**, 1715–1717.

Carrello, A., Ingley, E., Minchin, R. F., Tsai, S. & Ratajczak, T. (1999). *J. Biol. Chem.* **274**, 2682–2689.

Corpet, F. (1988). *Nucleic Acids Res.* **16**, 10881–10890.

Das, A. K., Cohen, P. T. W. & Barford, D. (1998). *EMBO J.* **17**, 1192–1199.

Duina, A. A., Kalton, H. M. & Gaber, R. F. (1998). *J. Biol. Chem.* **273**, 18974–18978.

Duina, A. A., Marsh, J. A. & Gaber, R. F. (1996). *Yeast*, **12**, 943–952.

Freeman, B. C., Toft, D. O. & Morimoto, R. I. (1996). *Science*, **274**, 1718–1720.

Galat, A. & Metcalfe, S. M. (1995). *Prog. Biophys. Mol. Biol.* (1995). **63**, 67–118.

Handschumacher, R. E., Harding, M. W., Rice, J. & Drugge, R. J. (1984). *Science*, **226**, 544–547.

Hoffmann, K. & Handschumacher, R. E. (1995). *Biochem. J.* **307**, 5–8.

Kallen, J., Spitzfalden, C., Zurini, M. G. M., Wider, G., Widmer, H., Wuthrich, K. & Walkinshaw, M. D. (1991). *Nature (London)*, **353**, 276–279.

Kieffer, L. J., Seng, T. W., Li, W., Osterman, D. G., Handschumacher, R. E. & Bayney, R. M. (1993). *J. Biol. Chem.* **268**, 12303–12310.

Kieffer, L. J., Thalhammer, T. & Handschumacher, R. E. (1992). *J. Biol. Chem.* **267**, 5503–5507.

Nair, S. C., Rimerman, R. A., Toran, E. J., Chen, S. Y., Prapapanich, V., Butts, R. N. & Smith, D. F. (1997). *Mol. Cell. Biol.* **17**, 594–603.

Pflugl, G., Kallen, J., Jansonius, J. M. & Walkinshaw, M. D. (1994). *J. Mol. Biol.* **244**, 385–409.

Pflugl, G., Kallen, J., Schirmer, T., Jansonius, J. N., Zurini, M. G. M. & Walkinshaw, M. D. (1993). *Nature (London)*, **361**, 91–94.

Pratt, W. B. & Toft, D. O. (1997). *Endocrine Rev.* **18**, 306–360.

Radanyi, C., Chambraud, B. & Baulieu, E. E. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11197–11201.

Ratajczak, T. & Carrello, A. (1996). *J. Biol. Chem.* **271**, 2961–2965.

Ratajczak, T., Carrello, A., Mark, P. J., Warner, B. J., Simpson, R. J., Moritz, R. L. & House, A. K. (1993). *J. Biol. Chem.* **268**, 13187–13192.

Rost, B., Sander, C. & Schneider, R. (1994). *Comput. Appl. Biosci.* **10**, 53–60.

Taylor, P., Page, A. P., Kondopidis, G., Husi, H. & Walkinshaw, M. D. (1998). *FEBS Lett.* **425**, 261–366.

Weisman, R., Creanor, J. & Fantes, P. (1996). *EMBO J.* **15**, 447–456.

Yokoyama, N., Hayashi, N., Seki, T., Pante, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, Y., Aebi, U., Fukui, M. & Nishimoto, T. (1995). *Nature (London)*, **376**, 184–188.

Young, J. C., Obermann, W. M. J. & Hartl, F. U. (1998). *J. Biol. Chem.* **273**, 18007–18010.

Structure, Vol. 9, 431–438, May, 2001, ©2001 Elsevier Science Ltd. All rights reserved. PII S0969-2126(01)00603-7

Two Structures of Cyclophilin 40: Folding and Fidelity in the TPR Domains

Paul Taylor,[1,5] Jacqueline Dornan,[1,5] Amerigo Carrello,[2,3] Rodney F. Minchin,[3] Thomas Ratajczak,[2,3] and Malcolm D. Walkinshaw[1,4]
[1] Structural Biochemistry Group
Institute of Cell and Molecular Biology
The University of Edinburgh
Michael Swann Building
King's Buildings
Mayfield Road
Edinburgh EH9 3JR
United Kingdom
[2] Department of Endocrinology and Diabetes
Sir Charles Gairdner Hospital
Hospital Avenue
[3] Department of Pharmacology
University of Western Australia
The Queen Elizabeth II Medical Centre
Nedlands WA 6009
Australia

Summary

Background: The "large immunophilin" family consists of domains of cyclophilin or FK506 binding protein linked to a tetratricopeptide (TPR) domain. They are intimately associated with steroid receptor complexes and bind to the C-terminal domain of Hsp90 via the TPR domain. The competitive binding of specific large immunophilins and other TPR-Hsp90 proteins provides a regulatory mechanism for Hsp90 chaperone activity.

Results: We have solved the X-ray structures of monoclinic and tetragonal forms of Cyp40. In the monoclinic form, the TPR domain consists of seven helices of variable length incorporating three TPR motifs, which provide a convincing binding surface for the Hsp90 C-terminal MEEVD sequence. The C-terminal residues of Cyp40 protrude out beyond the body of the TPR domain to form a charged helix—the putative calmodulin binding site. However, in the tetragonal form, two of the TPR helices have straightened out to form one extended helix, providing a dramatically different conformation of the molecule.

Conclusions: The X-ray structures are consistent with the role of Cyclophilin 40 as a multifunctional signaling protein involved in a variety of protein-protein interactions. The intermolecular helix-helix interactions in the tetragonal form mimic the intramolecular interactions found in the fully folded monoclinic form. These conserved intra- and intermolecular TPR-TPR interactions are illustrative of a high-fidelity recognition mechanism. The two structures also open up the possibility that partially folded forms of TPR may be important in domain swapping and protein recognition.

[4] Correspondence: m.walkinshaw@ed.ac.uk
[5] These authors contributed equally to this work.

Introduction

Cyclophilin 40 (Cyp40), FKBP52, and FKBP51 are the principal members of a family of large immunophilins found in all mammals. They comprise an N-terminal immunophilin domain with homology to the well-characterized single domains of cyclophilin or FK506 binding protein (FKBP) and an extended C-terminal region incorporating three units of the tetratricopeptide repeat (TPR) motif [1–5]. In *Saccharomyces cerevisiae*, two Cyp40 homologs, Cpr6 and Cpr7, encode proteins which bind the yeast transcriptional regulator and histone deacetylase Rpd3 [6]. Cpr7 has been observed in a trimeric complex with hsc82 and the Sti1 homolog Cns1. The Cpr7-Cns1 interaction is thought to be mediated via the cyclophilin domain of Cpr7 [7]. The detailed biological function of large immunophilins is not well understood; however, in all species they are associated with Hsp90 and play an important regulatory role in the maintenance of steroid receptor activity [8]. Within mature steroid receptor heterocomplexes, the large immunophilins compete directly for a common TPR acceptor site in Hsp90 [5, 9, 10].

The two structures described provide new insights into recognition mechanisms for TPR-containing proteins possibly using domain swapping and also suggest a folding pathway for the helical TPR domain.

Results and Discussion

The Cyclophilin Domain of Cyp40 Belongs to the Divergent Loop Family

Both monoclinic and tetragonal forms of Cyp40 (see Table 1) have very similar cyclophilin domains which resemble previously characterized single-domain cyclophilins [11, 12]. The "divergent loop" of bovine Cyp40 consists of an insert of 8 residues (60PTTGKPLH67). A consensus sequence for this additional loop, based on multiple sequence alignments, is **GK*LH [13]. A conserved glutamate (Glu96), which ties down the loop using an interesting hydrogen bonding motif, is found in a homologous position in most divergent loop cyclophilins (Figure 1). Like other divergent loop cyclophilins [13, 14], Cyp40 has two cysteines in close proximity in the reduced form. The side chains of Cys52 and Cys181 are well defined in the electron density and show a sulfur-sulfur distance of 5.4 Å. The formation of a disulphide bond in an oxidizing environment may induce a conformational change and provide a signaling mechanism in response to oxidative stress.

The peptidyl-prolyl *cis-trans* isomerase active site of Cyp40 is similar to that of human cyclophylin A (hCypA) [12]. The active site of the cyclophilin family is defined by the 13 amino acids found to be important in the binding of the drug cyclosporin A (CsA) to hCypA (Figure 1), and a number of these residues are highlighted in

Key words: Cyclophilin; Hsp90; immunophilin; tetratricopeptide repeat; TPR; protein recognition

Table 1. Crystal Parameters and Data Reduction Statistics

	Monoclinic Form	Tetragonal Form
λ	0.86	0.86
a	126.12	94.58
b	47.33	
c	85.56	118.33
β	119.42	
Space group	C2	$P4_22$
Unique reflections	40,561	33,904
Completeness (%)	98.3	90.5
D_{min}	1.8	2.0
$<I>/<\sigma(I)>$	20.67	19.6
R_{merge} (%)	5.0	5.7
R_{cryst}	17.83	24.32
R_{free}	25.63	30.79
Rmsd		
Bonds	0.07	0.06
Bond angle dists	0.23	0.22
Ramachandran plot		
N core region/N (%)	326/328 (99.4)	256/283 (97.3)
N inner core (%)	270/328 (82.3)	195/283 (74.1)

Figure 3a. The active site residues are identical, apart from His 141, which corresponds with Trp120 in hCypA (Figure 1). The PPIase activity of Cyp40 has a kcat/Km value of 1.9×10^6 $M^{-1}s^{-1}$ [15] compared with the value of 1.4×10^7 $M^{-1}s^{-1}$ for hCypA [16].

The TPR Domain of the Monoclinic Form Consists of Seven Helices and Is Attached to the Cyclophilin Domain by a Linker of 30 Amino Acids

The acidic linker region contains 11 aspartate and glutamate residues and forms a well-defined hydrogen-bonded structure with two β turns. The positively charged surface of helix Q forms salt bridges with linker residues D204 and D200 (D204/OD...K248/N = 2.99 Å, D200/OD2...K245/N = 2.73 Å, D200/OD1...K248/N = 3.16 Å; Figure 3a).

The seven helices labeled P, Q, R, S, T, U, and V form an extended helix-turn-helix pattern (Figure 3a). Each 34 amino acid TPR repeat consists of a (A_i, B_i) helix-turn-helix motif (Figure 2). The first three helices (P, Q, R) are longer than expected for idealized TPR sequences. Helix R incorporates a proline residue (P270), which also distorts the helical conformation of the 4 preceding residues. (A similar distortion occurs in all HEAT repeat helices which frequently contain proline [17].) The interhelix Ω angles [18] between adjacent helices are all positive (Ω[P, Q] = 13°, Ω[Q, R] = 26°, Ω[R, S] = 12°, Ω[S, T] = 24°, Ω[T, U] = 14°, Ω[U, V] = 22°). The structure of PP5 [19] has a similar arrangement of helices.

Figure 2 provides a comparison between the amino acid sequences of the C-terminal TPR-containing domains of large immunophilins with other Hsp 90 binding proteins, including PP5, Hop, and CNS1. A consensus sequence pattern of small and large hydrophobic residues has been postulated, with small residues occurring at positions 8, 20, and 27 [19]. The average A_i-B_i interhelix distance for each Cyp40 TPR is between 7.5 and 8 Å. This sharp turn brings the side chain of residue 8 close to the backbone of helix B and explains the requirement of the small Gly or Ala side chain at this position. The N-terminal domain (residues 18–181) of human Ser/Thr phosphatase PP5 [19] has an amino acid identity of 24% with Cyp40 over the corresponding 133 amino acids (Figure 2). A root mean square (rms) fit between the corresponding 133 Cα atoms is 1.2 Å

A Partially Folded TPR Domain Is Trapped in the Tetragonal Crystal Form

The structure of the tetragonal form retains only the first TPR motif (helices P and Q); the second TPR motif (helices R and S) has sprung out to form an elongated helix (Figure 3b). Residues comprising helices T, U, and V in the monoclinic form (Cα trace shown in Figure 4) are not visible in the electron density. This structure can therefore be regarded as a trapped intermediate in the folding pathway of the helix domain. It also suggests that conformations of such helical domains are flexible and that such helix extensions toward the end of a helical domain may occur in protein-protein recognition processes. Packing in the tetragonal lattice allows compensatory intermolecular helix-helix interactions (Figure 5a), and the intermolecular interactions between helices P and Q with helix R of an adjacent molecule are nearly identical to the intramolecular interactions found in the fully folded monoclinic form (Figure 5b). This provides an excellent example of domain swapping.

Helix V in the Monoclinic Form May Be Involved in Calmodulin Binding

Specific binding between large immunophilins and calmodulin has been shown in bovine Cyp40 [20], maize FKBP 66,[21], and rabbit FKBP52 [22]. Calmodulin binding requires an amphiphilic α helix that carries a net positive charge [23]. The calmodulin binding site to FKBP52 has been proposed to lie between residues 398–415 that corresponds to residues 353–370 of Cyp40. The X-ray structure of Cyp40 presented here shows that this putative interaction domain corresponds to the C-terminal portion of the V helix (Figure 3a), which is seen to extend beyond the shorter TPR helices. The protruding and predominantly positively charged helix V would therefore be accessible for calmodulin binding.

The X-Ray Structure of the Monoclinic Cyp40 TPR Domain Provides a Model Binding Surface for Hsp90

Site-directed mutagenesis of PP5 has shown that Lys97, Arg101, Lys32, and, to a lesser extent, Arg74 are required for Hsp90 binding [24]. A sequence alignment of TPR sequences (Figure 2) shows that residues Lys97 and Arg101 are conserved at positions 2 and 6 of TPR3 for all the related Hsp90 binding proteins. Deletion mutagenesis studies of the C-terminal domain of Hsp90 have shown that the conserved C-terminal sequence EEVD is important for immunophilin binding [25].

The monoclinic crystal form of Cyp40 grows exceptionally fast along the needle axis (needles grow to over 0.5 mm within minutes) and suggests a particularly energetically favorable interaction [26]. Analysis of the crystal packing (Figure 6) shows the C-terminal residues EKAAY fitting tightly into a groove on the surface formed by the three TPR motifs. The residues defining the

Figure 1. Alignment of Selected Cyclophilin Proteins

Sequence alignment of human cyclophilin A (CYPH_HUMAN), bovine and human cyclophilin 40 (CYP4_BOVIN and CYP4_HUMAN), WIS2 from *S. Pombe* (CYP4_SCHPO), and CPR6 and CPR7 from *S. cerevisiae* (CPR6_YEAST and CPR7_YEAST). The sequences were aligned using Clustal W [29]. In bold are the features discussed in the text, including conserved cysteines and glutamates and residues forming the inserted loop. The 13 residues important in defining the active site of the cyclophilin family of proteins are shown starred above the text. The high degree of conservation between these cyclophilin domains (residues 1–183) is clearly seen.

groove are Lys227, Phe234, Ser274, Asn278, Lys308, and Arg312. Significantly, the interaction involves close contacts with the three positively charged residues identified as critical for interaction with Hsp90 [24]. We have used this intermolecular interaction as a backbone template to model the Cyp40 interaction with the Hsp90 C-terminal MEEVD sequence (data not shown). This model requires the MEEVD peptide to run in the opposite direction to that found in the Hop-peptide complex [27], suggesting that the peptide binding groove of TPR domains may act as a peptide binding site with broad specificity.

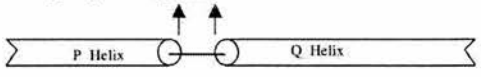
Biological Implications

The multiple but poorly understood roles of the large immunophilin family as peptidyl-prolyl isomerases, chaperones, and promiscuous binding proteins can be understood in part by the X-ray structure presented here. The classification of the cyclophilin domain as one of the divergent loop family implies that it plays a role in recognizing (as yet unidentified) partner proteins after phosphorylation or in signaling a response to oxidative stress. The two-domain picture of the Cyp40 molecule suggests a clear separation between the TPR binding domain docked onto the Hsp90 and the outward-facing enzyme active site, allowing independent functionality of the two domains. The MEEVD peptide binding groove in the TPR domain reveals a crucial recognition site for Hsp90-Cyp40 interaction. This Hsp-90 immunophilin interaction is competitive among the different large immunophilins (FKBP51, FKBP52, and Cyp40) that appear to modulate the interaction of Hsp90 with client proteins,

Alignment of TPR 1, TPR 2 and TPR 3 motifs of Hsp90 binding proteins.

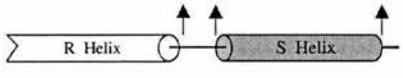
TPR 1 motifs of selected hsp90 binding proteins.

		10	20	30	34
CYP4_BOVIN	223-256	SEDLKNIGNT	FFKSQNWEMA	IKKYTKVLRY	VEGS
CYP4_HUMAN	223-256	TEDLKNIGNT	FFKSQNWEMA	IKKYAEVLRY	VDSS
FKBP5_HUMAN	268-301	AAIVKEKGTV	YFKGGKYMQA	VIQYGKIVSW	LEME
FKBP4_HUMAN	270-303	STIVKERGTV	YFKEGKYKQA	LLQYKKIVSW	LEYE
PPP5_HUMAN	28- 61	AEELKTQAND	YFKAKDYENA	IKFYSQAIEL	NPSN
CYP6_YEAST	219-252	IETVKNIGTE	QFKKQNYSVA	LEKYVKCDKF	LKEY
CYP7_YEAST	240-273	ANIIKESGTL	LFKKKDYSNA	FFKYRKSLNY	INEY
CNS1_YEAST	83-116	AENPKKQGNE	LYKAKRFKDA	RELYSKGLAV	ECED
IEFS_HUMAN	225-258	ALKEKELGND	AYKKKDFDTA	LKHYDKAKEL	DPTN



TPR 2 motifs of selected Hsp90 binding proteins.

		10	20	30	34
CYP4_BOVIN	273-306	LSCVLNIGAC	KLKMSDWQGA	VDSCLEALEI	DPSN
CYP4_HUMAN	273-306	LSCVLNIGAC	KLKMSNWQGA	IDSCLEALEL	DPSN
FKBP5_HUMAN	317-350	LAAFLNLAMC	YLKLREYTKA	VECCDKALGL	DSAN
FKBP4_HUMAN	319-352	LASHLNLAMC	HLKLQAFSAA	IESCNKALEL	DSNN
PPP5_HUMAN	62- 95	AIYYGNRSLA	YLRTECYGYA	LGDATRAIEL	DKKY
CYP6_YEAST	270-303	VSIPLNIAIC	ALKLKDYKQV	LVASSEVLYA	EAAD
CYP7_YEAST	292-325	MKIYLNLSLV	LFNLERYDDA	IMYATYLLEM	DNVP
CNS1_YEAST	121-154	ESLYANRAAC	ELELKNYRRC	IEDCSKALTI	NPKN
IEFS_HUMAN	300-333	MTYIINQAAV	YFEKGDYNKC	RELCEKAIEV	GREN



TPR 3 motifs of selected Hsp90 binding proteins.

		10	20	30	34
CYP4_BOVIN	307-340	TKALYRRAQG	WQGLKEYDQA	LADLKKAQEI	APED
CYP4_HUMAN	307-340	TKALYRRAQG	WQGLKEYDQA	LADLKKAQGI	APED
FKBP5_HUMAN	351-384	EKGLYRRGEA	QLLMNEFESA	KGDFEKVLEV	NPQN
FKBP4_HUMAN	353-386	EKGLFRRGEA	HLAVNDFELA	RADFQKVLQL	YPNN
PPP5_HUMAN	96-129	IKGYYRRAAS	NMALGKFRAA	LRDYETVVKV	KPHD
CYP6_YEAST	308-341	AKALYRRGLA	YYHVNDTDMA	LNDLEMATTF	QPND
CYP7_YEAST	330-363	AKAYYRRGNS	YLKKKRLDEA	LQDYIFCKEK	NPDD
CNS1_YEAST	155-188	VKCYYRTSKA	FFQLNKLEEA	KSAATFANQR	IDPE
IEFS_HUMAN	259-292	AKAYARIGNS	YFKEEKYKDA	IHFYNKSLAE	HRTP

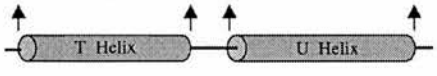


Figure 2. Alignment of TPR Sequences for Hsp90 Binding Proteins

Proteins that contain three copies only of the TPR motif and that have been shown to interact with Hsp90 were identified from the literature and are given in the table, using Swissprot names. Additionally, the TPR2 domain of Hop has been included. Each TPR motif from the above proteins was manually aligned. The proteins were bovine and human cyclophilin 40 (CYP4_BOVIN, CYP4_HUMAN); human FKBP 51 and 52 (FKBP5_HUMAN and FKBP4_HUMAN); human protein phosphatase 5 (PPP5_HUMAN); Cpr 6, Cpr 7, and Cns1 from *S. cerevisiae* (CYP6_YEAST, CYP7_YEAST, and CNS1_YEAST); and human Hop (IEFS_HUMAN). Numbering above the sequences runs from 1 to 34 and gives the position of a particular amino acid relative to the start of the TPR motif. The horizontal cylinders below each alignment designate the extent of helices P, Q, R, S, T, and U in the 34 amino acid sequence, using structural data from Cyp40 to assign helices. A broken cylinder indicates the helix begins or terminates earlier or later in the protein sequence, therefore failing to conform to the classical, idealized helix (13 aa) turn (3 aa) helix (14 aa) model for TPR motifs. Arrows projecting from the cylinders point to the positions in the sequence where the helices start and finish. The three linker residues between helix A_i and helix B_i are underlined. Closed and shaded cylinders illustrate a complete helix within the TPR motif. Only TPR 3 conforms to the idealized structure. Residues identified as being important in binding Hsp90 are bold and underlined. Structurally conserved residues are italicized and bold.

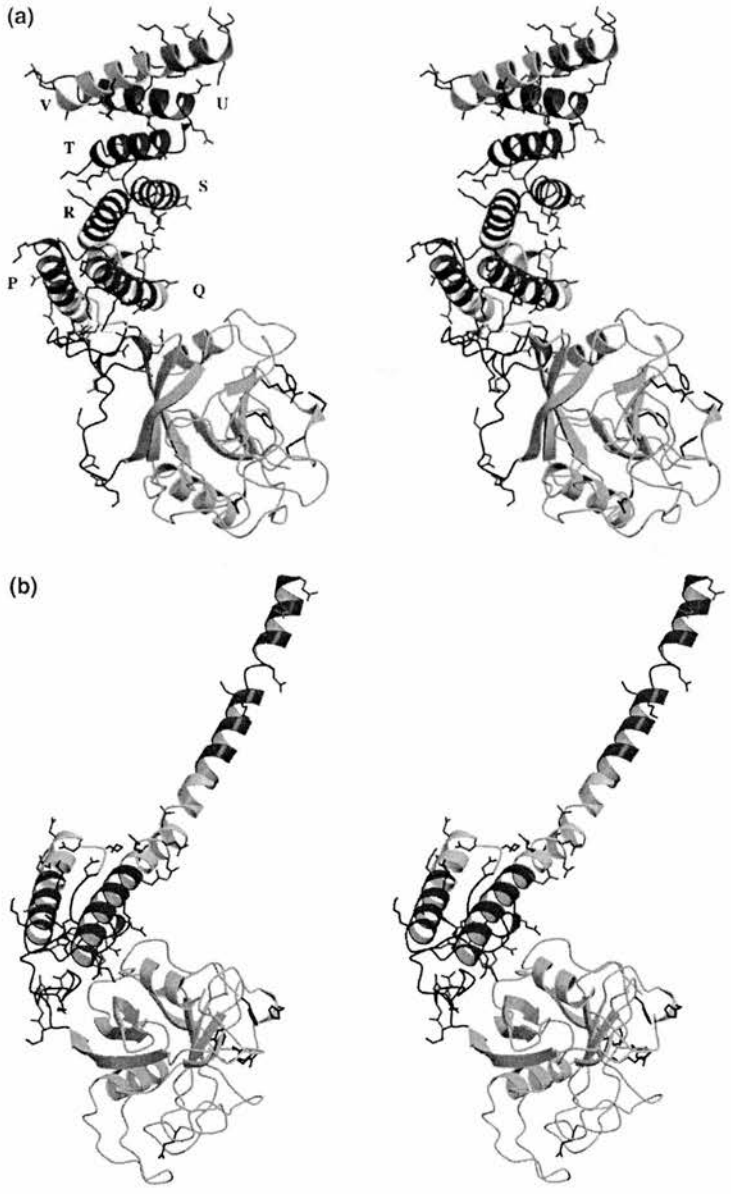


Figure 3. Stereo Pictures of the Monoclinic and the Tetragonal Folding Intermediate Forms of Bovine Cyclophilin 40

(a) The stereo picture of the monoclinic form of bovine cyclophilin 40 was drawn using Molscript [30]. The cyclophilin domain (residues 1–183) is colored green. Side chains of selected active site residues (Arg75, Phe80, Phe133, His141, His146) are colored black. Glu96, which forms hydrogen bonds to the divergent loop, is also shown. The backbone of the linker (residues 184–213) joining the cyclophilin and TPR domains is shown in black. Selected salt bridges between the positively charged residues of helix Q (K244, K245, K248, R251) with linker residues D204 and D200 (D204/OD...K248/N = 2.99 Å, D200/OD2...K245/N = 2.73 Å, D200/OD1...K248/N = 3.16 Å) are shown as dashed green lines. Side chains of all charged residues on the TPR domain are shown. Arg and Lys, blue; Glu and Asp, red. There are seven helices in the TPR domain: P (20 residues, V216-K235), Q (21 residues, W239-A259), R (24 residues, D262-K285), S (12 residues, W289-L300), T (13 residues, T307-G319), U (14 residues, Y323-I336), and V (18 residues, K341-K362). The slight distortion in helix R is caused by the presence of Pro270. The portions of the helices corresponding to the repeating TPR motifs shown in Figure 2 are colored dark blue (TPR1, helices P and Q, residues 223–256; TPR2, helices R and S, residues 273–306; TPR3, helices T and U, residues 307–340). The protruding C-terminal portion of helix V is the putative calmodulin domain and is colored magenta.

(b) The stereo picture of the tetragonal folding intermediate form of bovine cyclophilin 40 was drawn using Molscript [30]. Coloring is the same as (a). Helices T, U, and V are not visible in the electron density. Helices are colored according to the fully folded monoclinic form in (a): Helix P (20 residues, V216-K235), Q (21 residues, W239-A259), R (24 residues, D262-K285), S (12 residues, W289-D298). The portions of the helices corresponding to the repeating TPR motifs shown in Figure 2 are colored dark blue (TPR1, residues 223–256; TPR2, residues 273–298). Alternative conformations of residues Ala260 and Ser287 are primarily responsible for helices R and S adopting an extended rod-like helical conformation.

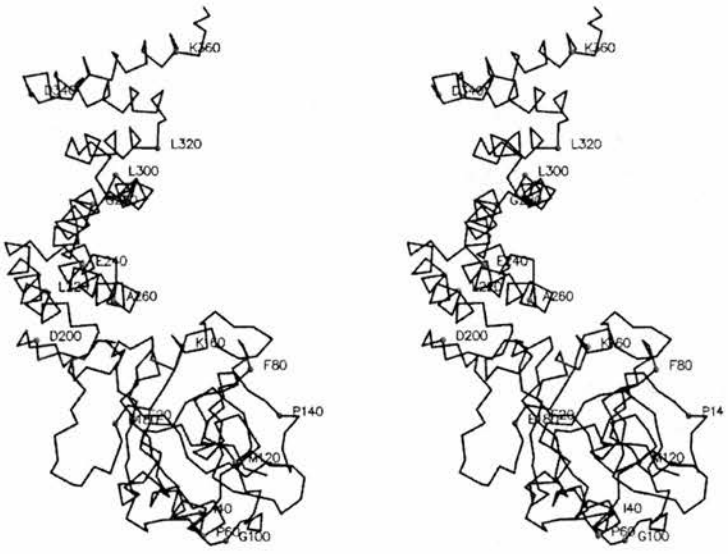


Figure 4. Stereo Cα Trace of the Monoclinic Form of Bovine Cyclophilin Amino Acids Are Labeled Every 20th Residue

including steroid receptors and kinases. The two Cyp40 structures presented here provide the basis for detailed site-directed mutation studies to map out the complete binding surfaces mediating the Hsp90-immunophilin recognition. Once bound to Hsp90, it is likely that the large immunophilins act to anchor or trigger conformational changes in the Hsp90-bound client proteins and may therefore play a key role in Hsp90-regulated signal transduction pathways. Inhibition of the Hsp90-immunophilin interaction by small drug molecules could provide a new target for a wide range of therapeutic areas, including cancer and inflammation.

Experimental Procedures

Crystallization

The hanging drop vapor diffusion method was used to grow crystals in 24-well Linbro plates at 4°C.

Monoclinic Crystals

In most drops, very long thin monoclinic needles (2 × 0.01 × 0.005 mm) grow within hours. Occasionally, much larger plates formed with dimensions up to 0.6 × 0.2 × 0.15 mm. The 1 ml well solution was composed of 320 μl of a 50% solution of PEG 4000, 100 μl of 1 M MES buffer (pH 6.1), 200 μl of 50% v/v glycerol solution, 100 μl 4 M NaCl, and 280 μl water. The 4 μl hanging drop consisted of 2 μl well solution plus 2 μl of protein at 10–80 mg/ml in 20 mM Tris (pH 8) buffer containing 100 mM NaCl.

Tetragonal Crystals

The well solution consisted of 480 μl MPD, 100 μl 1 M imidazole at pH 6.5, 140 μl 50% v/v glycerol solution, and 280 μl water. The hanging drop consisted of 2 μl well solution plus 2 μl of protein at 10–80 mg/ml in 20 mM Tris (pH 8) buffer containing 100 mM NaCl. Teardrop-shaped crystals with dimensions up to 0.2 × 0.15 × 0.15 mm grew within a few days.

Structure Determination

Data used in the final refinement for both structures were collected on station 9.6 at the Daresbury synchrotron.

Monoclinic Structure

The complete model of Cyp3 from *C. elegans* [13] was used for a molecular replacement search using AMORE [28]. This model gave a single 10 σ peak in the rotation function, and this rotated model gave a 9 σ peak in the translation function. The initial model gave an R factor of 55%, and, from the first map, it was possible to change the residues in the cyclophilin domain to match the Cyp40 sequence. Additionally, some residues of the TPR domain, mainly within helices, were visible, and, over many cycles of model fitting and refinement, it was possible to trace almost the entire TPR domain. The final model contains 489 waters. No ions were found; however, one glycerol molecule was located. The final refined structure runs from Ser2 to Tyr365. The 5 C-terminal residues were not visible in the electron density.

Tetragonal Form

Initial stages of the structure determination were similar to the monoclinic form with clear rotation and translation solutions apparent. The final model contains 224 waters and one glycerol molecule. The refined structure runs from Ser2 to Asp298. The C-terminal 72 residues were not visible in the electron density.

Acknowledgments

We thank Morag Bilsland for a detailed sequence analysis and for helpful discussions. We thank the Wellcome Trust for support and the staff at the Synchrotron Radiation Source (SRS) (Daresbury) and the European Synchrotron Radiation Facility (ESRF) (Grenoble) for help with synchrotron data collection.

Received: October 25, 2000
Revised: March 13, 2001
Accepted: April 13, 2001

References

1. Lebeau, M.C., et al., and Baulieu, E.E. (1992). P59, an hsp 90-binding protein. Cloning and sequencing of its cDNA and preparation of a peptide-directed polyclonal antibody. J. Biol. Chem. *267*, 4281–4284.
2. Peattie, D.A., et al., and Benasutti, M. (1992). Expression and characterization of human fkbp52, an immunophilin that associates with the 90-kda heat-shock protein and is a component of steroid-receptor complexes. Proc. Natl. Acad. Sci. USA *89*, 10974–10978.
3. Kieffer, L.J., Seng, T.W., Li, W., Osterman, D.G., Handschumacher, R.E., and Bayney, R.M. (1993). Cyclophilin-40, a protein with homology to the p59 component of the steroid-receptor

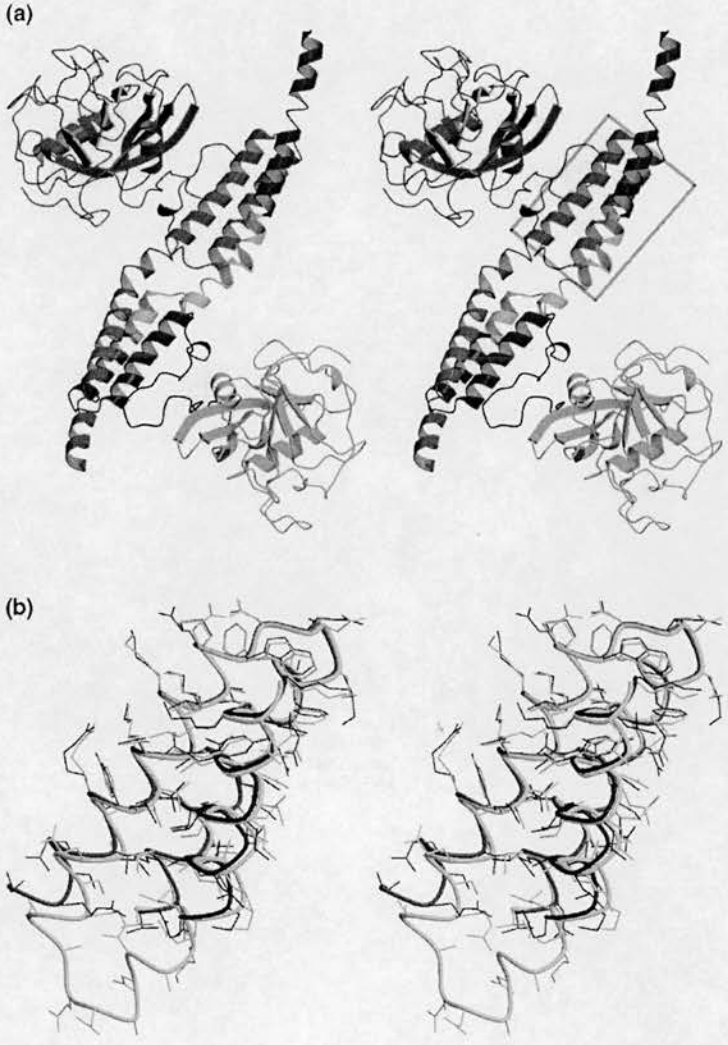


Figure 5. Stereo Picture of the Main Intermolecular Interaction in the Tetragonal Form of Cyp40 and an Overlay of the Folding Intermediate Dimer and the Fully Folded Conformer

(a) The partially unfolded TPR domains form a symmetrical dimer. One molecule is colored as in Figure 3b, and the partner molecule is colored red. The first TPR, comprising helices P and Q (shown in the boxed region), make an intermolecular contact with helix R′ of the dimer-related molecule.

(b) The boxed region of (b), with the same color scheme, showing helices P, Q, and R′ for the tetragonal form, overlaid with helices P, Q, and R (yellow) of the folded TPR domain of the monoclinic form. The intra- and intermolecular interactions of helix R with helices P and Q are seen to be nearly identical.

complex. Cloning of the cDNA and further characterization. J. Biol. Chem. *268*, 12303–12310.

4. Ratajczak, T., et al., and House, A.K. (1993). The cyclophilin component of the unactivated estrogen-receptor contains a tetratricopeptide repeat domain and shares identity with p59 (fkbp59). J. Biol. Chem. *268*, 13187–13192.
5. Nair, S.C., et al., and Smith, D.F. (1997). Molecular cloning of human fkbp51 and comparisons of immunophilin interactions with hsp90 and progesterone receptor. Mol. Cell. Biol. *17*, 594–603.
6. Duina, A.A., Marsh, J.A., and Gaber, R.F. (1996). Identification of 2 cyp-40-like cyclophilins in saccharomyces cerevisiae, one of which is required for normal growth. Yeast *12*, 943–952.
7. Dolinski, K.J., Cardenas, M.E., and Heitman, J. (1998). Cns1 encodes an essential p60/sti1 homolog in saccharomyces cerevisiae that suppresses cyclophilin 40 mutations and interacts with hsp90. Mol. Cell. Biol. *18*, 7344–7352.
8. Duina, A.A., Chang, H.C.J., Marsh, J.A., Lindquist, S., and Gaber, R.F. (1996). A cyclophilin function in hsp90-dependent signal-transduction. Science *274*, 1713–1715.
9. Owens-grillo, J.K., et al., and Pratt, W.B. (1995). The cyclosporine a-binding immunophilin CyP-40 and the FK506-binding immunophilin hsp56 bind to a common site on hsp90 and exist in independent cytosolic heterocomplexes with the untransformed glucocorticoid receptor. J. Biol. Chem. *270*, 20479–20484.
10. Ratajczak, T., and Carrello, A. (1996). Cyclophilin-40 (cyp-40), mapping of its hsp90 binding domain and evidence that fkbp52 competes with cyp-40 for hsp90 binding. J. Biol. Chem. *271*, 2961–2965.
11. Braun, W., Kallen, J., Mikol, V., Walkinshaw, M.D., and Wuthrich, K. (1995). Three-dimensional structure and actions of immunosuppressants and their immunophilins. FASEB J. *9*, 63–72.
12. Kallen, J., Mikol, V., Taylor, P., and Walkinshaw, M.D. (1998). X-ray structures and analysis of 11 cyclosporin derivatives complexed with cyclophilin a. J. Mol. Biol. *283*, 435–449.
13. Dornan, J., et al., and Walkinshaw, M.D. (1999). Biochemical and structural characterization of a divergent loop cyclophilin from Caenorhabitis elegans. J. Biol. Chem. *274*, 34877–34883.
14. Taylor, P., Page, A.P., Kontopidis, G., Husi, H., and Walkinshaw, M.D. (1998). The x-ray structure of a divergent cyclophilin from the nematode parasite brugia malayi. FEBS Lett. *425*, 361–366.
15. Kieffer, L.J., Thalhammer, T., and Handschumacher, R.E. (1992). Isolation and characterization of a 40-kda cyclophilin-related protein. J. Biol. Chem. *267*, 5503–5507.
16. Liu, J., Albers, M.W., Chen, C.M., Schreiber, S.L., and Walsh, C.T. (1990). Cloning, expression, and purification of human cyclophilin in escherichia-coli and assessment of the catalytic

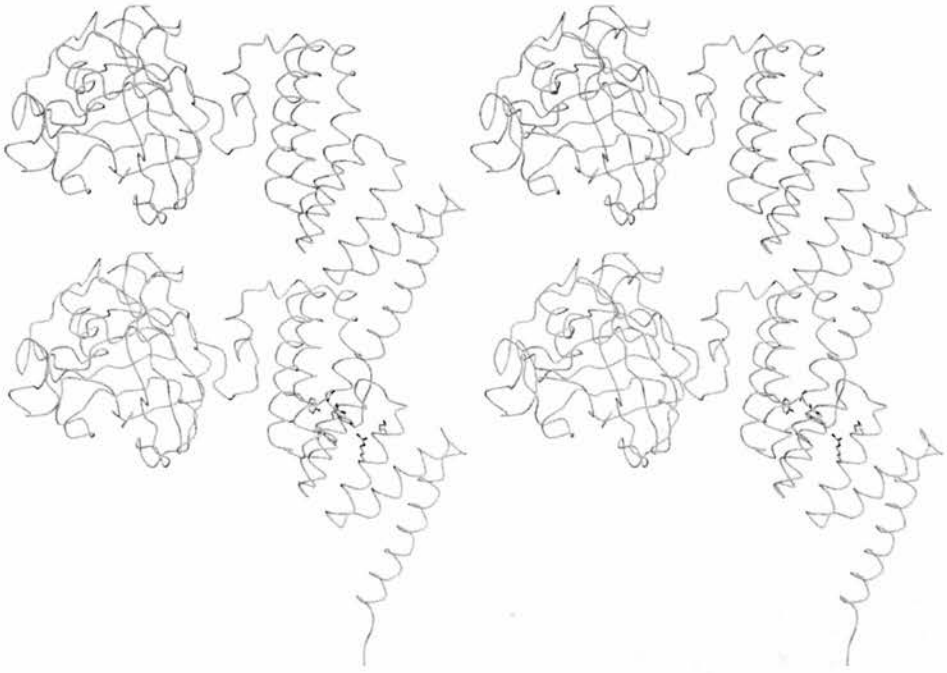


Figure 6. Stereo Picture of the Main Intermolecular Interaction in the Monoclinic Form of Cyp40

The yellow and blue molecules are related by a unit cell translation along b of 47.3 Å. The C-terminal residues EKAAY from one molecule (blue) fit into a groove of an adjacent molecule (yellow). The groove is formed by 2 residues from each of the three TPR helices ($A_{1,2,3}$): Lys227, Phe234, Ser274, Asn278, Lys308, and Arg312 (side chains shown). Tyr365 fits into a hydrophobic pocket in the groove and is the last residue visible in the electron density. (The C-terminal 5 residues are presumably structurally disordered in the crystal.)

role of cysteines by site-directed mutagenesis. Proc. Natl. Acad. Sci. USA *87*, 2304–2308.

17. Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A., and Barford, D. (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. Cell *96*, 99–110.
18. Bowie, J.U. (1997). Helix packing angle preferences. Nat. Struct. Biol. *4*, 915–917.
19. Das, A.K., Cohen, P.T.W., and Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for tpr-mediated protein-protein interactions. EMBO J. *17*, 1192–1199.
20. Ratajczak, T., Carrello, A., and Minchin, R.F. (1995). Biochemical and calmodulin-binding properties of estrogen-receptor binding cyclophilin expressed in escherichia-coli. Biochem. Biophys. Res. Commun. *209*, 117–125.
21. Hueros, G., Rahfeld, J., Salamini, F., and Thompson, R. (1998). A maize FK506-sensitive immunophilin, mzFKBP-66, is a peptidylproline cis-trans-isomerase that interacts with calmodulin and a 36-kDa cytoplasmic protein. Planta *205*, 121–131.
22. Massol, N., Lebeau, M.C., Renoir, J.M., Faber, L.E., and Baulieu, E.E. (1992). Rabbit fkbp59-heat shock protein-binding immunophillin (hbi) is a calmodulin binding-protein. Biochem. Biophys. Res. Commun. *187*, 1330–1335.
23. Ghirlanda, G., Lear, J.D., Lombardi, A., and DeGrado, W.F. (1998). From synthetic coiled coils to functional proteins: Automated design of a receptor for the calmodulin-binding domain of calcineurin. J. Mol. Biol. *281*, 379–391.
24. Russell, L.C., Whitt, S.R., Chen, M.S., and Chinkers, M. (1999). Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. J. Biol. Chem. *274*, 20060–20063.
25. Carrello, A., Ingley, E., Minchin, R.F., Tsai, S., and Ratajczak, T. (1999). The common tetratricopeptide repeat acceptor site for steroid receptor-associated immunophilins and Hop is located in the dimerization domain of hsp90. J. Biol. Chem. *274*, 2682–2689.
26. Dornan, J., Taylor, P., Carrello, A., Minchin, R.F., Ratajczak, T., and Walkinshaw, M.D. (1999). Purification characterisation and crystallisation in two ctrystal forms of bovine Cyclophilin 40. Acta Crystallogr. D Biol. Crystallogr. 1079–1082.
27. Scheufler, C., et al., and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell *101*, 199–210.
28. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. *A 50*, 157–163.
29. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.
30. Kraulis, P.J. (1991).Molscript: a program to produce both detailed and schematic plots of protein structures. J. Applied Crystallogr. *24*, 946–950.

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of ocr, the product of gene 0.3 of bacteriophage T7

Shane S. Sturrock,[a] David T. F. Dryden,[b] Constandache Atanasiu,[a] Jacqueline Dornan,[a] Sandra Bruce,[a] Andrew Cronshaw,[a] Paul Taylor[a] and Malcolm D. Walkinshaw[a]*

[a]Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, Scotland, and [b]Department of Chemistry, University of Edinburgh, Edinburgh, Scotland

Correspondence e-mail: m.walkinshaw@ed.ac.uk

Received 22 May 2001
Accepted 10 July 2001

Ocr, the product of gene 0.3 of bacteriophage T7, prevents the action of restriction endonucleases of the host bacteria. The amino-acid sequence of ocr has less than 20% similarity to any protein of known three-dimensional structure. Ocr has been crystallized in a number of different crystal forms and X-ray data for the seleno-L-methionine-substituted form has been collected to a resolution of 1.8 Å. The presence of caesium was found to be required for good crystal growth. Anomalous X-ray data was used to identify possible positions for Se and Cs atoms in the unit cell.

1. Introduction

Bacterial cells have evolved a wide range of defence mechanisms against viral infection. One such mechanism is effected through restriction-modification (R-M) systems such as the type I R-M enzymes (Bickle & Kruger, 1993; Murray, 2000). These enzymes allow the bacteria to recognize their own DNA through the strain-specific methylation of bases by transferring methyl groups from *S*-adenosyl-methionine to specific adenine or cytosine bases in the recognition site. Where the host DNA is hemimethylated (only one strand is modified) after DNA replication, the enzyme will methylate the other strand to produce fully methylated DNA. Cleavage of unmethylated DNA (for example, from an invading bacteriophage) by type I R-M systems requires two type I R-M enzymes attached at two distant recognition sites. A complex DNA translocation process driven by ATP hydrolysis occurs prior to DNA cleavage (Studier & Bandyopadhyay, 1988; Ellis *et al.*, 1999; Berge *et al.*, 2000).

To counteract the bacterial R-M systems, bacteriophages have developed anti-restriction systems (Kruger & Schroeder, 1981; Kruger & Bickle, 1983; Bickle & Kruger, 1993). It has been shown that bacteriophage T7 is not subject to restriction or modification by *Escherichia coli* B and *E. coli* K type I R-M systems *in vivo* (Eskridge *et al.*, 1967) but is *in vitro* (Eskin *et al.*, 1973; Bandyopadhyay *et al.*, 1985). The ocr protein product of gene 0.3, a 26 kDa dimer (Mark & Studier, 1981), is responsible for this protection (Studier, 1975). *Ocr* is the first gene on T7 DNA and is the first expressed after infection (Studier, 1972, 1973). Ocr is thought to act as a polyanion mimic of DNA and competes with DNA for the DNA-binding site of the type I R-M enzymes (Dunn *et al.*, 1981; Bandyopadhyay *et al.*, 1985). The inhibition may be similar to the interaction of uracil glycosylase inhibitor with the DNA repair enzyme uracil glycosylase (Mol *et al.*, 1995; Savva & Pearl, 1995; Putnam *et al.*, 1999). Ocr is very acidic: 34 of its amino acids are either aspartic or glutamic acids and only six are arginine or lysine (Dunn *et al.*, 1981) (Fig. 1). Structure predictions (Dryden *et al.*, 1995; Sturrock & Dryden, 1997) have presented a working model of the architecture of the binding site for ocr on a type I R-M enzyme. A crystal structure will provide a clear insight into how ocr mimics DNA and the nature of the interaction of ocr with type I R-M enzymes.

2. Results and discussion

2.1. Protein production and characterization

The wild-type ocr protein was produced by transformation of *E. coli* strain BL21(DE3)-pLysS with the plasmid pAR2993. Transformed cells were plated out on agar with 20 μg ml^{-1} ampicillin and 25 μg ml^{-1} chloramphenicol to produce single colonies. A single colony was used to innoculate flasks containing 500 ml LB media supplemented with ampicillin and chloramphenicol. When the optical density at 650 nm reached 0.6, IPTG was added to a final concentration of 0.4 m*M*. After incubation for 2 h at 310 K, the cells were harvested by centrifugation. The cells were sonicated and cell debris removed by centrifugation for 3 h at 30 000*g*. The supernatant was loaded onto a 200 × 16 mm DEAE Sepharose fast flow ion-exchange column equilibrated with 20 m*M* Tris–HCl, 7 m*M* β-mercaptoethanol, 300 m*M* NH_4Cl at pH 8 at 48 ml h^{-1}. A 500 ml gradient from 0.3 to 1 *M* NH_4Cl was run overnight at 24 ml h^{-1}. Frac-

Table 1
Data sets.

λ (Å)	0.87	1.2	0.9777	0.9793
Max. resolution	1.8	1.85	1.8	1.8
Completeness (final shell)	99.3 (98.5)	97 (85)	99.1 (96.7)	99.1 (97.1)
R_{merge} (final shell)	4.9 (9.6)	4.8 (12)	5.7 (13.5)	4.7 (15.9)
Total reflections	71916	65144	81164	72034
Unique	8927	8922	9012	9033
Anomalous diffraction ratio (%)	4.98	5.20	6.51	5.19
Se (f', f'')	−1.0, 3.0	−1.4, 0.73	−6.9, 0.5	−6.9, 0.5
Cs (f', f'')	0.0, 3.0	0.0, 5.3	0.2, 3.7	0.2, 3.7

† Anomalous diffraction ratio = $\langle||F+| - |F-||\rangle/\langle 1/2(|F+| + |F-|)\rangle$.

Table 2
Fractional coordinates of the Se and Cs atoms used to generate the Patterson peaks shown in Fig. 3.

	x	*y*	*z*
Se1	0.08063	0.27275	0.24952
Se2	0.20312	0.16840	0.51414
Se3	0.08258	−0.00231	0.04150
Se4	0.10748	−0.31247	0.00709
Cs1	0.15214	−0.41763	0.18218
Cs1	0.08887	0.12099	−0.11881
Cs1	0.33431	0.20409	0.09952

tions containing ocr were put through a HiLoad 16/60 Superdex 200 preparative grade gel-filtration column in 20 m*M* Tris–HCl, 7 m*M* β-mercaptoethanol, 300 m*M* NH_4Cl pH 8 buffer. An additional step was tried using a 150 × 16 mm phenyl Sepharose affinity column equilibrated in 20 m*M* Tris–HCl, 7 m*M* β-mercaptoethanol, 50 m*M* NH_4Cl, 2.4 *M* $(NH_4)_2SO_4$ buffer at pH 8. Protein was applied to the column and flushed with this buffer for 30 min at 60 ml h^{-1} and protein was then eluted at 48 ml h^{-1} with a 500 ml gradient from 2.4 to 0 *M* $(NH_4)_2SO_4$. This achieved a higher level of purity but was not found to be necessary for the production of good crystals. Chromatography columns were from Amersham Pharmacia. The purified ocr protein was stored in 20 m*M* Tris–HCl, 7 m*M* β-mercaptoethanol, 300 m*M* NH_4Cl pH 8 buffer containing 50% glycerol at 253 K.

A selenomethionine-substituted form of ocr was produced by transforming the methionine-deficient *E. coli* strain B834. This strain (kindly provided by Dr W. Hunter, Dundee University, Scotland) was transformed with pAR2993 as with BL21(DE3) and grown in minimal media supplemented with 20 μg ml^{-1} of all amino acids except cysteine (to prevent reutilization of sulfur to make normal methionine) and with 20 μg ml^{-1} seleno-L-methionine along with 20 μg ml^{-1} ampicillin.

Electrospray mass spectra were measured using a Themoquest LCQ mass spectrometer (Thermoquest Corporation, Paradise, Hemel Hempstead, England). Native ocr gave M_r = 13 677 ± 1 Da, compared with a mass of M_r = 13 959 ± 1 Da for the selenium form. The difference in mass of 47 Da × 6 indicates 100% incorporation of six Se atoms.

2.2. Crystallization and X-ray data analysis

Crystallization was carried out using the hanging-drop vapour-diffusion method. The drop consisted of equal volumes (normally 2 μl) of protein solution and well solution. Crystallization trials were performed with Hampton Research sets I and II and crystalline precipitate was observed with polyethylene glycol (PEG) 8000, Tris–HCl buffer and sodium acetate. Refinement of these conditions led to well formed crystals in 32% PEG, 300 m*M* NH_4Cl, 0.1 *M* Tris–HCl pH 8.0 and 0.2 *M* sodium acetate at 291 K for both native and selenomethionine forms of the protein (type I). This crystal form was metastable and after about 10 d the crystals gradually dissolved. In some of the drops, poorly formed plate-like crystals reappeared after days or weeks (type II).

Caesium acetate was also used in crystal trials and found to produce large well formed crystals of type I. Trials with rubidium acetate and potassium acetate yielded progressively poorer crystals ranging between the caesium and sodium acetate crystal quality. No crystals were obtained when ammonium acetate was used as an additive. Crystals of type I, irrespective of size or external morphology, diffracted poorly. An initial data set for the wild type was collected to 2.5 Å in space group *C*2, with unit-cell parameters a = 88.18, b = 105.96, c = 73.34 Å, β = 120.319°.

The observation that caesium seemed to help crystal growth was further investigated. Using CsCl, a large plate-like crystal (Fig. 2) grew within two weeks under the following conditions: well solution 32% PEG 8000, 0.1 *M* Tris–HCl pH 8.2, 0.2 *M* CsCl, 2.8% glycerol; drop 2 μl well solution plus 2 μl of protein solution at 16 mg ml^{-1}. This selenomethionine form of the ocr crystals could be flash-frozen in liquid nitrogen without use of additional cryoprotectant. These crystals also belong to space group *C*2 and have unit-cell parameters a = 78.78, b = 37.66, c = 35.49 Å, β = 98.41°. Anomalous data sets were collected for the selenomethionine crystals on station 9.5 at the Daresbury synchrotron and station BM30 at the ESRF, Grenoble (Table 1).

The fluorescence spectrum was measured to determine the absorption edge of selenium and data sets were measured at four wavelengths; the inflection point of the edge (0.9777 Å, λ_2), the peak of the white-line absorption (0.9793 Å, λ_3) and at two remote wavelengths (0.87 Å, λ_1; 1.2 Å, λ_4). The merging R factor between Friedel pairs within each data set does indicate a slightly larger anomalous signal at the absorption edge (6.9% compared with the remote R factors of 5.7) and this is supported by the size of the anomalous diffraction ratios for the four data sets (Table 1). Anomalous Patterson difference maps for each of the four data sets were calculated (Fig. 3). Major differences between these maps is explained by the unexpected discovery that the crystals incorporated significant amounts of caesium, which has a large anomalous contribution over this wavelength range (Fig. 4).

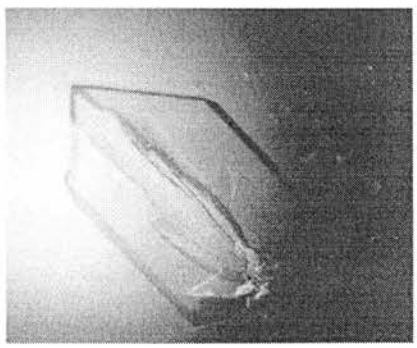


Figure 2
A crystal of the selenium form of ocr with dimensions 0.4 × 0.4 × 0.1 mm.

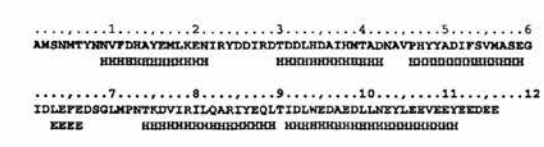


Figure 1
Sequence of ocr. The secondary-structure prediction for ocr is from the program *PHD* (Rost & Sander, 1993).

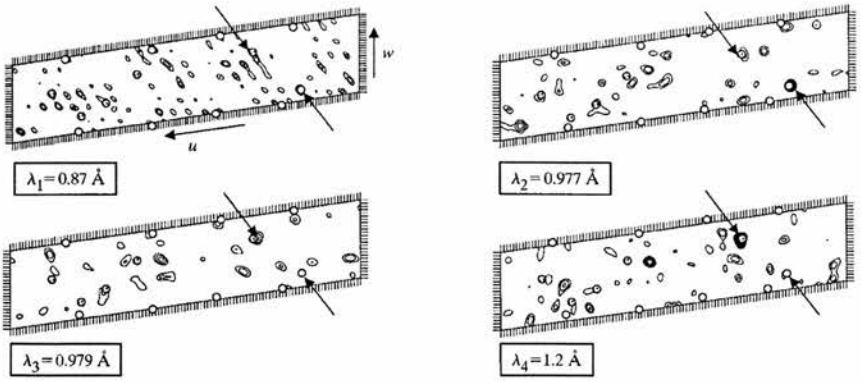


Figure 3
Harker sections calculated at $y = 0$. The Fourier coefficients for calculating the Patterson map for each of the four data sets are $(|F_+| - |F_-|)^2$. The section runs from $u = 0$ to 1 in 108ths and from $w = 0$ to 0.5 in 48ths. The stippled circles correspond to the Harker peaks resulting from the Cs—Cs vectors from three independent caesium ions. The pale filled circles correspond to the Harker peaks of the Se—Se vectors from four of the Se atoms. Patterson maps were calculated and plotted using the program *CNS* (Brunger *et al.*, 1998). Arrows indicate positions of the strongest selenium and caesium peaks.

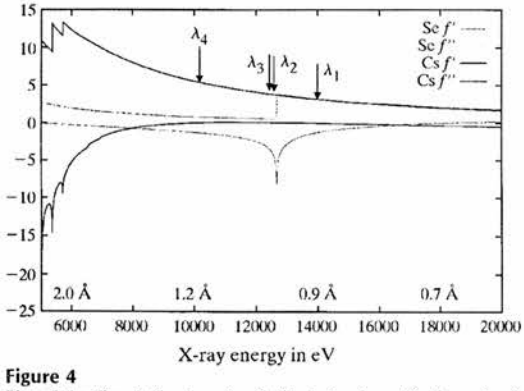


Figure 4
Plot of the f' and f'' values for Cs (dark lines) and Se (gray lines) from 6000 eV (2 Å) to 18 000 eV (0.7 Å) and plotted using tools from the web page http://www.bmsc.washington.edu/scatter.

A good solution to the Patterson maps has been obtained using the program *SOLVE* and most of the Harker peaks can be explained by four Se atoms and three Cs atoms. The strongest selenium peak (u = 0.17, 0, 0.08) is in the map using data calculated at the absorption edge λ_2 = 0.98 Å (Fig. 3). This Patterson peak is also strong in the map calculated from data measured at λ_1 = 0.87 Å. The selenium signal is very much reduced at the longer wavelengths above the selenium absorption edge. The anomalous Patterson peaks calculated with data measured at λ_3 = 0.981 Å and λ_4 = 1.2 Å are dominated by the caesium signal. This is clearly seen by the presence of the strong peak u = 0.31, v = 0, w = 0.38 (Fig. 3). The changes in intensities of the anomalous Patterson peaks can be explained by inspection of Fig. 4, which shows that the anomalous contribution of caesium varies between 6 electrons and 4 electrons over the four wavelengths, while the selenium contribution is only significant at the shorter wavelengths. The position of the heavy atoms (Table 2) determined by *SOLVE* (Terwilliger & Berendzen, 1999) has provided an electron-density map that is of sufficient quality to solve the structure. Preliminary refinement has confirmed the assignment of the selenium and caesium positions.

We thank the Leverhulme Trust for providing funds for this research work (Grant F/158/BC) and the staffs of ESRF and Daresbury SRS for help with data collection. DD thanks the Royal Society for a University Research Fellowship and CA was supported by a Darwin Trust studentship. Plasmid pAR1993 was a kind gift from William Studier and Alan Rosenberg (Brookhaven National Laboratories).

References

Bandyopadhyay, P. K., Studier, F. W., Hamilton, D. L. & Yuan, R. (1985). *J. Mol. Biol.* **182**, 567–578.
Berge, T., Ellis, D. J., Dryden, D. T. F., Edwardson, J. M. & Henderson, R. M. (2000). *Biophys. J.* **79**, 479–484.
Bickle, T. A. & Kruger, D. H. (1993). *Microbiol. Rev.* **57**, 434–450.
Brunger, A. T., Adams, P. A., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
Dryden, D. T. F., Sturrock, S. S. & Winter, M. (1995). *Nature Struct. Biol.* **2**, 632–635.
Dunn, J. J., Elzinga, M., Mark, K.-K. & Studier, F. W. (1981). *J. Biol. Chem.* **256**:2579–2585.
Ellis, D. J., Dryden, D. T., Berge, T., Edwardson, J. M. & Henderson, R. M. (1999). *Nature Struct. Biol.* **6**, 15–17.
Eskin, B., Lautenberger, J. A. & Linn, S. (1973). *J. Virol.* **11**, 1020–1023.
Eskridge, R., Weinfield, H. & Paigen, K. (1967). *J. Bacteriol.* **93**, 835–844.
Kruger, D. H. & Bickle, T. A. (1983). *Microbiol. Rev.* **47**, 345–360.
Kruger, D. H. & Schroeder, C. (1981). *Microbiol. Rev.* **45**, 9–51.
Mark, K.-K. & Studier, F. W. (1981). *J. Biol. Chem.* **256**, 2573–2578.
Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W. & Tainer, J. A. (1995). *Cell*, **82**, 701–708.
Murray, N. E. (2000). *Microbiol. Mol. Biol. Rev.* **64**, 412–434.
Putnam, C. D., Shroyer, M. J., Lundquist, A. J., Mol, C. D., Arvai, A. S., Mosbaugh, D. W. & Tainer, J. A. (1999). *J. Mol. Biol.* **282**, 331–346.
Rost, B. & Sander, C. (1993). *J. Mol. Biol.* **232**, 584–599.
Savva, R. & Pearl, L. H. (1995). *Nature Struct. Biol.* **2**, 752–757.
Studier, F. W. (1972). *Science*, **176**, 367–376.
Studier, F. W. (1973). *J. Mol. Biol.* **79**, 237–248.
Studier, F. W. (1975). *J. Mol. Biol.* **94**, 283–295.
Studier, F. W. & Bandyopadhyay, P. K. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 4677–4681.
Sturrock, S. S. & Dryden, D. T. F. (1997). *Nucleic Acids Res.* **25**, 3408–3414.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.